



Magisterarbeit

Neue Chancen und Möglichkeiten für  
Ordnungssystematiken durch Virtualisierung:  
Anwendung am Beispiel der Erfassung und Klassifizierung  
des informationswissenschaftlichen Bücherbestandes der  
Universitätsbibliothek Hildesheim

Peter Hanke

E-Mail: [hanke\\_peter@web.de](mailto:hanke_peter@web.de)

Universität Hildesheim

Institut für Angewandte Sprachwissenschaft

Studiengang: Internationales Informationsmanagement

Schwerpunkt: Angewandte Informationswissenschaft

Erstgutachterin: Prof. Dr. Christa Womser-Hacker

Zweitgutachter: Dr. Thomas Mandl

Hildesheim, Mai 2002

## **Kurzfassung des Inhaltes der Arbeit**

Die vorliegende Arbeit beschäftigt sich mit der Erfassung und Klassifizierung informationswissenschaftlicher Titel in der Universitätsbibliothek Hildesheim. Da für diese Bücher bisher keine eigene Klassifikation existierte und sie unter vielen unterschiedlichen Systemstellen abgelegt sind, soll den Studierenden durch ein „virtuelles Bibliotheksregal“ ein leichter Zugriff auf Bücher aus der Informationswissenschaft ermöglicht werden.

## **Thesis Abstract**

This thesis provides a classificatory scheme for Information Science books to be implemented by the library of the University of Hildesheim. This „virtual shelf“ will allow students ready access to books which are not grouped together on shelves as no classification for this discipline and related topics is in place yet.

# Inhaltsverzeichnis

<b>0</b>	<b>Einleitung.....</b>	<b>1</b>
<b>1</b>	<b>Allgemeine Aspekte von Ordnungssystematiken in Bibliotheken ....</b>	<b>3</b>
1.1	Die Notwendigkeit von Bibliothekskatalogen .....	3
1.2	Die Funktionen von Katalogen.....	5
1.2.1	Bestandserschließung in Katalogen .....	6
1.2.1.1	Die formale Erschließung .....	6
1.2.1.2	Die inhaltliche Erschließung .....	7
1.2.2	Standortkataloge .....	8
1.2.3	Online-Kataloge.....	9
1.3	Bibliographische Klassifikationssysteme.....	11
1.3.1	Die Dewey Decimal Classification (DDC) .....	12
1.3.2	Die Universal Decimal Classification (UDC) .....	15
1.3.3	Die Library of Congress Classification (LCC) .....	16
1.3.4	Die Regensburger Verbundklassifikation (RVK) .....	18
1.3.5	Die Basisklassifikation (BK).....	19
<b>2</b>	<b>Die Problematik der informationswissenschaftlichen Literaturrecherche in der Universitätsbibliothek Hildesheim .....</b>	<b>21</b>
2.1	Verdeutlichung der Problematik anhand von Literaturbeispielen .....	22
2.2	Vergleiche mit anderen Bibliotheken mit eigener informationswissenschaftlicher Klassifikation .....	27
2.2.1	Auswahl von geeigneten Bibliotheken bzw. Klassifikationen.....	27
2.2.2	Vergleich mit der Situation Studierender in Konstanz, Regensburg und Saarbrücken .....	31
<b>3</b>	<b>Evaluierung ausgewählter Bibliotheken mit dem Sachgebiet Informationswissenschaft.....</b>	<b>34</b>
3.1	Auswertung der Literaturrecherche.....	36
3.1.1	Gesamtausstattung .....	36
3.1.2	Ausstattung der einzelnen Themengebiete .....	37
3.1.3	Erfassung der Titel in den einzelnen Klassifikationen.....	41
3.1.3.1	Saarbrücken.....	41
3.1.3.2	Regensburg.....	43
3.1.3.3	Konstanz .....	45
3.1.3.4	Thematische Kongruenzen zwischen den Klassifikationen .....	49
3.2	Evaluierung der informationswissenschaftlichen Klassifikationen.....	50

<b>4</b>	<b>Der informationswissenschaftliche Bücherbestand in der UB Hildesheim.....</b>	<b>57</b>
4.1	Alphabetisches Register von informationswissenschaftlichen Begriffen und zugehörigen Notationen.....	57
4.2	Erfassung der relevanten Titelmenge in der UB Hildesheim über den Web-OPAC .....	59
4.3	Abgleich mit dem informationswissenschaftlichen Bestand der UB Konstanz.....	62
4.3.1	Das Parsen der Literaturbestände.....	64
4.3.1.1	Die Programmiersprache gawk .....	64
4.3.1.2	Parsen und Aufbereiten des kid .....	66
4.3.1.3	Parsen und Aufbereiten des Hildesheimer Bestandes .....	70
4.3.2	Der Abgleich des Konstanzer und Hildesheimer Bestandes.....	75
4.4	Literaturanschaffungen im Institut für Angewandte Sprachwissenschaft.....	78
4.5	Suchanfragen an den Web-OPAC der UB Hildesheim .....	79
<b>5</b>	<b>Konzept einer Zugriffsmöglichkeit auf informationswissen- schaftliche Bücher über eine virtuelle Ordnungssystematik.....</b>	<b>80</b>
5.1	Erarbeitung einer informationswissenschaftlichen Klassifikation .....	81
5.2	Einordnung des Hildesheimer Bestandes in die neue Klassifikation.....	86
5.2.1	Zuordnung der Hildesheimer Notationen zu der Klassifikation .....	86
5.2.2	Zuweisung der gefundenen Titel auf die Klassifikation .....	86
5.3	Zugriff über ein HTML-Dokument.....	87
5.4	Verlinkung einzelner Systemstellen der neuen Klassifikation mit Sys- temstellen anderer informationswissenschaftlicher Klassifikationen .....	89
<b>6</b>	<b>Evaluierung .....</b>	<b>90</b>
<b>7</b>	<b>Ausblick .....</b>	<b>98</b>
	<b>Literaturverzeichnis.....</b>	<b>103</b>
	<b>Anhang .....</b>	<b>107</b>

## Abbildungsverzeichnis

Abbildung 1: Regensburger Verbundklassifikation (RVK), Informationswissenschaft .....	19
Abbildung 2: Trefferanzeige Web-OPAC UB Hildesheim, Volker Walter .....	25
Abbildung 3: Trefferanzeige Web-OPAC UB Hildesheim, Rainer Kuhlen.....	26
Abbildung 4: Trefferanzeige Web-OPAC UB Hildesheim, Peter Ludes .....	27
Abbildung 5: Beispiel für eine Trefferausgabe im Konstanzer Koala .....	46
Abbildung 6: Ergebnis der Suchanfrage „Dokument“ an die Notationsbeschreibungen der UB Hildesheim .....	58
Abbildung 7: Beispiel des Ergebnisses einer Suchanfrage an den OPC4 .....	61
Abbildung 8: Auszug aus der ersten Konstanzer Bestandsdatei (Titel mit kid Signaturen) .....	66
Abbildung 9: Auszug aus der geparsten Konstanzer Bestandsdatei des kid (Signaturen) .....	68
Abbildung 10: Auszug aus der zweiten Konstanzer Bestandsdatei (Titel mit kid- Signaturen („Ex.“) und kid-Systemstelle („Not.“)) .....	69
Abbildung 11: Auszug aus der Hildesheimer Bestandsdatei .....	71
Abbildung 12: Fünf Beispiele für die Verwendung des jeweils erst genannten Autors der Hildesheimer Bestandsdatei im kid .....	73
Abbildung 13: Ausschnitt aus der bearbeiteten Hildesheimer Bestandsdatei (Autor: Kurztitel   Signatur   Nebenstellen) .....	75
Abbildung 14: HTML-Seite „Information und Gesellschaft / Informationspolitik“	88

## Tabellenverzeichnis

Tabelle 1: Recherche nach Hildesheimer Titeln über verschiedene Notationen .....	23
Tabelle 2: Erfassung der Hildesheimer Titel in Klassifikationen der anderen Standorte .....	32
Tabelle 3: Treffermengen der informationswissenschaftlichen Literaturrecherche in vier Bibliotheken.....	38
Tabelle 4: Einordnung recherchierter Saarbrücker Titel in der dortigen Klassifikation.....	42
Tabelle 5: Einordnung recherchierter Regensburger Titel in der dortigen Klassifikation.....	44
Tabelle 6: Einordnung recherchierter Regensburger Titel in der dortigen Klassifikation.....	48
Tabelle 7: Vergleich der Einordnung von 16 Titeln in der Konstanzer und Hildesheimer Klassifikation .....	92
Tabelle 8: Titel aus OPAC-Suchanfragen in der Hildesheimer Klassifikation ...	94

## 0 Einleitung

In den letzten Jahren hat die Digitalisierung von Bibliotheken zunehmend an Bedeutung gewonnen. Für Universitätsbibliotheken ist es heute selbstverständlich, dass sie einen OPAC (Online Public Access Catalogue) anbieten. Auf diese Weise soll zum einen den Studierenden vor Ort ein möglichst leichter Zugriff auf die vorhandenen Dokumente gestattet werden, zum anderen soll das Angebot auch für einen größeren Benutzerkreis über das Internet verfügbar gemacht werden.

Onlinekataloge sind konventionellen Katalogen als Erschließungs- und Nachweisinstrumente weit überlegen. Durch die Möglichkeit zur Virtualisierung der Inhalte gewinnt der digitale Katalog an Flexibilität und Leistungsfähigkeit. Der Mehrwert für den Benutzer ist enorm. Allerdings kann er nur dann in vollem Umfang von den zusätzlichen Angeboten profitieren, wenn die Bibliotheksinhalte in einer Weise aufbereitet sind, die ihm einen guten Überblick über den Bestand erlauben. Auch die Qualität digitaler Bibliotheken hängt ganz entscheidend von der Güte der Klassifikationen ab.

Einen konkreten Anwendungsfall für die virtuelle Aufbereitung bibliothekarischer Inhalte stellen die informationswissenschaftlichen Titel der Universitätsbibliothek Hildesheim dar. Das Fach verfügt im Rahmen der Hildesheimer Aufstellungssystematik über keine eigene Systemstelle, so dass aufgrund der Interdisziplinarität dieses Faches die Literatur über viele Sachgebiete verstreut steht. Im Gegensatz zu den Kommilitonen anderer Fachbereiche fehlt den Studierenden der Angewandten Informationswissenschaft (als Schwerpunkt im Hildesheimer Studiengang „Internationales Informationsmanagement“ angeboten) eine geeignete Repräsentation der geführten Titel.

Die vorliegende Arbeit möchte den Hildesheimer Studierenden der Informationswissenschaft bessere Möglichkeiten der Literaturrecherche offerieren; die zur Verfügung stehende Zeit hatte zur Folge, dass sich dieses Vorhaben auf den Bücherbestand beschränkt. Nach theoretischen Überlegungen zu bibliothekarischen Ordnungssystematiken (Kapitel 1) werden die angesprochenen

Probleme bei der Suche nach informationswissenschaftlicher Literatur verdeutlicht (Kapitel 2). Anschließend folgt eine genauere Begutachtung anderer Universitätsbibliotheken mit einer eigenen informationswissenschaftlichen Klassifikation (Kapitel 3). Die Erfassung des in der Hildesheimer UB vorliegenden Bestandes an informationswissenschaftlichen Titeln (Kapitel 4) geht der Erarbeitung einer Zugriffsmöglichkeit auf diese Literatur (Kapitel 5) voraus. Eine kurze Bewertung des Erreichten (Kapitel 6) und ein Ausblick auf notwendige Folgearbeiten (Kapitel 7) schließen sich an. Die einzelnen Kapitel stellen dabei keinen chronologischen Ablauf dar, sondern bauen inhaltlich aufeinander auf.

Hinweise auf die verwendete Literatur finden sich direkt in Klammern im Text. Um durch lange URLs nicht die Lesbarkeit zu erschweren, wurden hier stellvertretend die Organisationsbezeichnungen gewählt. Mehrere Seiten derselben Organisation werden durch hochgestellte Zahlen voneinander unterschieden (z.B. OCLC Forest Press<sup>2</sup>).

An dieser Stelle soll auch Frau Maria Elisabeth Müller, Leiterin der UB Hildesheim, und Herrn Edwin Pretz, EDV-Beauftragter und Fachreferent für Informatik, Naturwissenschaften und Technik in der UB Hildesheim, für ihre Hilfe und Zusammenarbeit gedankt werden. In allen Phasen der Arbeit zeigten sie sich sehr entgegenkommend und waren für die Beantwortung entstehender Fragen jederzeit ansprechbar. Mein Dank gilt ebenfalls Herrn Gerhard Schmitz-Veltin, in der Universitätsbibliothek Konstanz Fachreferent für Physik und Mathematik und zuständig für die dortige informationswissenschaftliche Klassifikation „kid“, welcher mir insbesondere beim Abgleich der Bücherbestände geholfen hat.



# **1 Allgemeine Aspekte von Ordnungssystematiken in Bibliotheken**

## **1.1 Die Notwendigkeit von Bibliothekskatalogen**

Eine Bibliothek wird ganz wesentlich durch die Größe und die Ordnung ihres Bestandes charakterisiert. Erst die Erschließung des Bestandes macht eine Ansammlung von Büchern zu einer Bibliothek. Dabei gilt es, eine Ordnungssystematik festzulegen, die geeignet ist, alle Bücher sinnvoll aufzunehmen.

Historisch gesehen, war die Aufstellung des Bibliotheksbestandes nach einer zweckmäßigen Systematik bis ins späte 18. Jahrhundert die wichtigste bibliothekarische Aufgabe (vgl. Haller 1998, 16). Die Aufstellung bewirkt eine bestimmte Abfolge von Büchern und anderen Materialien, wie z.B. Zeitschriften und CD-ROMs. Sie werden lokal zusammengefasst und in einer thematischen Reihenfolge eingeordnet.

Als zweite Aufgabe von Bibliotheken schloss sich später die Erstellung von Katalogen an. Um die Bedeutung eines Katalogs im bibliothekarischen Sinn auszuführen, kann man die Definition der Anglo-American Cataloguing Rules (AACR-2) heranziehen:

„1. A list of materials contained in a collection, a library, or a group of libraries, arranged according to some definite plan. 2. In a wider sense, a list of materials prepared for an particular purpose, e.g., an exhibition catalogue, a sales catalogue“ (zit. n. Haller 1998, 17).

Die Notwendigkeit der Erfassung des Bestandes in einem Katalog ergibt sich, wenn

- Bücher und Medien unsystematisch aufgestellt sind,
- der Bestand einen gewissen Umfang überschreitet,
- eine von der Art der Aufstellung unabhängige Übersicht erreicht werden soll,
- Informationen geliefert werden sollen, die den Bestand einer einzelnen Bibliothek erfassen aber auch darüber hinaus gehen.

(vgl. Haller 1998, 16)

Jede Bibliothek verfügt über einen Standortkatalog, der die Bücher in der Reihenfolge ihrer Aufstellung verzeichnet. Da dahinter ein inhaltlicher Plan steht, ist der Standortkatalog ein Standort gebundener Systematischer Katalog, also ein Sachkatalog (vgl. Haller 1998, 17). Weitere wichtige Arten von Katalogen stellen der Schlagwortkatalog, der Alphabetische Katalog und der Online-Katalog dar.

Der Vorgang der Schlagwortkatalog-Erstellung setzt sich im wesentlichen aus einer Inhaltsanalyse des Dokumentes mit einer entsprechenden Begriffsfindung, einer Terminologiekontrolle und einer regelgerechten Vergabe der Schlagworte zusammen (vgl. Haller 1998, 205). Das Schlagwort dient einer vom Titel unabhängigen, inhaltlichen Erschließung des Dokumentes. Die zu benutzenden Begriffe werden in einer Reihe von Nachschlagewerken verbindlich festgelegt. In erster Linie werden deutsche Terme verwendet, es sei denn, ein nicht deutscher Begriff ist aussagekräftiger.

Der Alphabetische Katalog führt den gesamten Bestand an Dokumenten in alphabetischer Ordnung auf, und zwar geordnet nach dem Anfangsbuchstaben des Verfassers bzw. dem Sachtitel eines Werkes. Er soll bei der Suche nach einem Werk helfen, von dem der Verfasser bzw. der Titel bekannt sind.

In der Form des Online-Kataloges OPAC (Online Public Access Catalogue) ist der Bibliothekskatalog eine Datenbank mit strukturierten Datensätzen für Bücher und andere Bibliotheksmaterialien. Die Ordnung wirkt sich hier in der Art und der Form der Suchbegriffe aus, die nach bestimmten Regeln den Datensätzen als Attribute zugeordnet werden. Auf diese Weise wird über elektronische Netze der Zugriff auf Kataloge fremder Bibliotheken und deren Dokumente möglich. Dadurch entsteht eine Form von virtueller Bibliothek, die ein weit größeres Angebot an Informationen bereithält, als es eine einzelne Bibliothek jemals könnte.

Aufstellungen und Kataloge ergänzen sich gegenseitig in ihren Funktionen. Eine Aufstellung ermöglicht die Suche im Rahmen der Aufstellungssystematik. Es kann direkt geprüft werden, ob das gesuchte Buch den Vorstellungen ent-

spricht. Entliehene Bücher allerdings sind am Standort nicht verfügbar, und jedes Buch kann nur unter einem Aspekt aufgestellt werden (vgl. Haller 1998, 18).

Kataloge ermöglichen sowohl eine formale, als auch eine sachliche Suche. Die formale Suche erlaubt Fragen nach Elementen des Titels, wie z.B. dem Namen des Verfassers, dem Sachtitel, dem Herausgeber, der Signatur und dem Erscheinungsdatum. Bei der sachlichen Suche wird nach inhaltlichen Angaben geforscht. Ein indirektes Prüfen anhand der bibliographischen Beschreibung kann Auskunft über die mögliche Verwendung des Dokumentes geben. Auch entliehene Bücher sind aufgeführt und ein Buch kann mehrmals in verschiedenen Rubriken erschlossen werden. Soweit es sich um einen Online-Katalog mit digitalisierten Dokumenten handelt, kann der Text unmittelbar eingesehen werden (vgl. Haller 1998, 18).

Mithilfe der Ordnungssystematik von Katalogen wird jedem Buch oder Dokument ein bestimmter Standort mit einer Signatur in der Bibliothek zugewiesen. Ein einzelnes Buch kann außerdem unter verschiedenen Aspekten aufgefunden werden (vgl. Haller 1998, 18). Kataloge sind somit zur Benutzung und Verwaltung einer Bibliothek unabdingbar.

### **1.2 Die Funktionen von Katalogen**

Anfangs waren Bibliothekskataloge ausschließlich als Handwerkszeug für Bibliothekare gedacht. Die ersten Publikums kataloge gibt es erst seit Ende des 19. Jahrhunderts (vgl. Haller 1998, 19). Heute bietet sich die Technik des elektronischen Kataloges an, um für Benutzer und Bibliothekare völlig neue und umfassendere Möglichkeiten der Ordnung und Beschaffung von Literatur zu eröffnen.

### **1.2.1 Bestandserschließung in Katalogen**

Kataloge bilden die Basis der Bestandserschließung. Sie sollen eine schnelle und zuverlässige Auskunft über den Bestand einer Bibliothek geben. Man unterscheidet dabei zwei grundsätzliche Formen der Erschließung:

- die vom Titel ausgehende formale Erschließung (Formalerschließung)
- die vom Inhalt ausgehende inhaltliche Erschließung (Inhaltsererschließung oder Sacherschließung)

(vgl. Haller 1998, 20)

#### *1.2.1.1 Die formale Erschließung*

Diese Art der Erschließung, die auch als Formalkatalogisierung (descriptive cataloguing) bezeichnet wird (vgl. Haller 1998, 20), orientiert sich an allen Daten, die zum Titel des Dokumentes gehören. In dieser Erfassungsart werden beispielsweise

- der Titel,
- der Name des Autors,
- der Verfasser,
- der Verlag,
- der Erscheinungsort,
- das Erscheinungsjahr,
- die ISBN (International Standard Book Number),
- die ISSN (International Standard Serial Number),
- die ISMN (International Standard Music Number),
- der Dokumententyp (Dissertation, Enzyklopädie, Festschrift, Landkarte, Notendruck, Schriftenreihe, Wörterbuch, Zeitschrift, usw.),
- und der Medientyp (Buch, Audiovisuelles Material, CD, CD-ROM, Film, Spiel, usw.) aufgenommen.

(vgl. Haller 1998, 40)

Ein Problem dabei ist, dass der Titel einerseits eine inhaltliche Beschreibung sein soll, andererseits aber auch Interesse wecken soll, um so als Kaufanreiz zu dienen. Aus diesem Grund ist zumindest der Hauptsachtitel in vielen Fällen

nur bedingt als Information bei der Suche nach Literatur zu einem bestimmten Themengebiet zu gebrauchen.

### *1.2.1.2 Die inhaltliche Erschließung*

Bei dieser Erschließungsform von Materialien, auch Sachkatalogisierung (subject cataloguing) genannt, wird der Inhalt unabhängig von den Titeldaten analysiert und erfasst. Bei der bibliothekarischen Sacherschließung unterscheidet man zwei verschiedene Arten:

- die verbale Inhaltserschließung, basierend auf natürlich-sprachlichen Bezeichnungen (Schlagwörter und Deskriptoren)
- die klassifikatorische Inhaltserschließung, beruhend auf den Notationen hierarchischer Klassifikationssysteme

(vgl. Haller 1998, 20)

Wird bei der Suche auf diese Ordnungsmechanismen zurückgegriffen, so weiß der Suchende in der Regel nicht, ob für seine Fragestellung, sein Problem, ein geeignetes Dokument existiert (vgl. Haller 1998, 41). Man spricht in diesem Fall von einer „Problemrecherche“. Dabei muss der Benutzer mit der Art und Form der Katalogisierung vertraut sein und Erfahrungen bei dieser Art der Suche besitzen, um Erfolg zu haben.

Um das richtige Schlagwort zu verwenden, ist es von Bedeutung, dass der User grundlegende Kenntnisse des bei der Verschlagwortung benutzten Vokabulars hat. Im Vordergrund steht die punktuelle Suche nach bestimmten Themen. Die Verschlagwortung erfolgt in den meisten Bibliotheken nach den RSWK (Regeln für den Schlagwortkatalog), die auch in Bibliotheken in Österreich und der Schweiz Anwendung finden. Für jeden Sachverhalt wird das begrifflich engste Schlagwort gesucht, also z.B. Jagd- oder Königsschloss statt nur Schloss (vgl. Haller 1998, 205). Nach ihrer Art werden fünf Hauptkategorien unterschieden:

- Personenschlagwörter (Personennamen)
- Geographische und ethnographische Schlagwörter (z.B. USA, Dänisch)
- Sachschlagwörter (z.B. Erkenntnistheorie)
- Zeitschlagwörter (z.B. Geschichte 1848)

- Formschlagwörter (Art und Form der Darstellung, z.B. Zeitschrift, Wörterbuch)

(vgl. Haller 1998, 205 ff)

Eine weitere Schwierigkeit ist, dass nur etwa 50 Prozent der Dokumente in Hochschulbibliotheken mit Schlagwörtern erschlossen werden. Bis zu 80 Prozent dieser Dokumente sind deutsch- oder englischsprachig, so dass auch hier eine deutliche Einschränkung bei den Recherchemöglichkeiten gegeben ist (vgl. Haller 1998, 41). Vorteilhaft wirkt sich hingegen aus, dass in der sogenannten Schlagwortnormdatei (SWD) Homonyme und Synonyme der Schlagwörter assoziativ und hierarchisch verknüpft sind. Auf diese Weise wird automatisch nach verwandten Begriffen gesucht.

Bei der Suche in den Notationen (als Synonym wird auch der Begriff „Systemstelle“ verwendet) einer Klassifikation steht dem Benutzer eine hierarchische Ordnung zur Verfügung, die ihm das ganze Umfeld eines Fachgebietes in Zusammenhang stellt. Um so übersichtlicher, logischer und verständlicher die Ordnungssystematik ist, um so leichter wird es dem Suchenden gemacht, an sein Ziel zu kommen. Unter diesem Gesichtspunkt sollte im Rahmen der vorliegenden Arbeit für den Bereich der Informationswissenschaft eine Klassifikation geschaffen werden, die dem Benutzer die Suche nach fachspezifischer Literatur erleichtert.

Eine weitere Form von Suchbegriffen neben den formalen und sachlichen stellen im Online-Katalog die Stichwörter dar, die sowohl der formalen als auch der inhaltlichen Erschließung dienen. Im Gegensatz zum genormten Vokabular von Schlagwörtern sind hier alle Formen sprachlicher Ausdrucksmöglichkeiten zulässig.

### **1.2.2 Standortkataloge**

Der Standortkatalog vereinigt in der Regel in sich die Aspekte des Formal- und des Sachkataloges. Zum einen werden bei der systematischen Aufstellung die Inhalte als Grundlage für die Ordnung herangezogen, zum anderen werden die

Bücher innerhalb eines Bereichs nach formalen Angaben sortiert. Die wichtigsten formalen Aufstellungsarten sind die alphabetische (beispielsweise nach dem Verfassernamen), die chronologische (beispielsweise nach dem Bearbeitungsdatum) und die nach der physischen Form (beispielsweise Printmedium, elektronisches Dokument) (vgl. Haller 1998, 252).

Im Standortkatalog sind die Bücher und Medien vollständig in der Reihenfolge ihrer Aufstellung nachgewiesen. Da er somit den Gesamtbestand nachweist, spricht man auch von einem Hauptkatalog (vgl. Haller 1998, 65). Er ist entscheidend für die Vergabe der Signaturen, wodurch der Standort der Dokumente festgelegt wird. Bei systematischer Aufstellung kann er als Sachkatalog dienen. Außerdem hat er die Aufgabe eines Inventarverzeichnisses (vgl. Haller 1998, 252).

Aus der Funktion als Sachkatalog, die von der Qualität seiner inhaltlich-systematischen Aufstellung abhängt, ergeben sich besondere Problematiken. Beispielsweise können sogenannte Doppelstellen auftreten. Dabei handelt es sich um gleiche Sachverhalte, die an mehreren Stellen der Systematik auftreten. Das kann der Fall sein bei räumlich getrennten Unterbringungen, bei Fächerüberschneidungen, bei besonderen Forschungsschwerpunkten der Universität oder bei Handapparaten zu speziellen Themen in Seminaren (vgl. Haller 1998, 257). Außerdem kann bei der systematischen Aufstellung nur jeweils ein inhaltlicher Aspekt eines Buches berücksichtigt werden.

In diesen Fällen kann eine virtuelle Ordnungssystematik die Probleme beheben. Die Bücher können dann unter verschiedenen Notationen an beliebig vielen Stellen der Systematik geführt werden, wobei nur eine Notation signaturbildend wirkt.

### **1.2.3 Online-Kataloge**

Ein Online-Katalog bzw. virtueller Katalog erlaubt dem Benutzer eine größere Freiheit bei der Materialsuche. Er kann einzelne Themengebiete durchforsten, ohne befürchten zu müssen, dass ihm ausgeliehene Werke entgehen. Er ge-

winnt einen schnellen Überblick über das gesamte Angebot einer Bibliothek in einem bestimmten Teilbereich. Neben einer gezielten Suche nach Literatur kann der Benutzer sich auch ziellos und unsystematisch durch den Katalog bewegen. Dabei trifft er unter Umständen auf Dokumente, von deren Existenz er nichts ahnte.

Über einzelne Suchbegriffe oder Kombinationen von mehreren Suchtermen kann man schnell viele relevante Dokumente aus einem Sachgebiet finden. Dabei sind auch Kombinationen von formaler und inhaltlicher Suche möglich, also beispielsweise eine Suchanfrage bestehend aus dem Verfassernamen und einem Schlagwort. Neben dieser Erweiterung der Recherchemöglichkeiten beinhalten Online-Kataloge die Problematik der eingeschränkten Fehlertoleranz. Suchbegriffe und Indexeinträge werden genau miteinander abgeglichen. Auch ist es, wie schon erwähnt, für eine exakte Schlagwortsuche erforderlich, zumindest grundlegende Kenntnisse des benutzten Thesaurus zu besitzen.

Um die Suche möglichst erfolgreich zu gestalten, wird deshalb versucht, Elemente einer Ähnlichkeitssuche und der Vergleich von Wortphrasen mit einzubauen. Dabei müssen einerseits die Suchbegriffe in den Suchindizes und andererseits die Sucheingaben maschinell entsprechend aufbereitet werden (vgl. Haller 1998, 51). Suchbegriff und Indexbegriff müssen einer Zeichenbehandlung unterzogen werden, um so das Suchergebnis möglichst unabhängig von der Schreibweise zu machen.

Trotzdem ergeben sich insbesondere bei der Suche nach Abkürzungen in der Form von Initialbuchstaben Schwierigkeiten. Die Problematik liegt hier im Weglassen oder Hinzufügen von Leerschritten (Spatien) bei der Aufnahme eines Titels in den Index. Wird beispielsweise der Buchtitel „K. u. K. Monarchie“ in der Form „K.u.K. Monarchie“ (also ohne Spatien) indexiert, so werden zwar sowohl die Suchbegriffe „K.u.K. Monarchie“ als auch „KuK Monarchie“ gefunden. Nicht erfolgreich sein wird allerdings die Suche nach „K. u. K. Monarchie“ und „K u K Monarchie“, obwohl das Buch eventuell in der Bibliothek verfügbar ist (vgl. Hal-



ler 1998, 55f). Bei der Suche im Web-OPAC der UB Hildesheim lässt sich diese Aussage am genannten Beispiel verifizieren.

Man sieht, welche Schwierigkeiten sich bei einer Materialsuche mit formalen Suchangaben bzw. Schlagwörtern ergeben können. Durch die Benutzung einer geeigneten Klassifikation können diese aus der Schreibweise resultierenden Probleme vermieden werden. Einige wissenschaftliche Disziplinen, wie beispielsweise die Informationswissenschaft, sind aufgrund ihres jüngeren Entstehungsdatums nur unzureichend oder gar nicht in Klassifikationen abgebildet. Ein sogenannter „virtueller Katalog“ kann hier Abhilfe schaffen, indem unabhängig von der Aufstellungssystematik der Bibliothek eine neue Ordnungssystematik erstellt wird, die den Bestand unter anderen Kriterien erfasst. Diese Magisterarbeit soll, bezogen auf den informationswissenschaftlichen Bücherbestand der Universitätsbibliothek Hildesheim, hierzu einen Beitrag leisten und ein virtuelles Bücherregal realisieren.

### **1.3 Bibliographische Klassifikationssysteme**

Wie bereits erwähnt, besteht ein Teil der inhaltlichen Erschließung in der Zuordnung in hierarchische Klassifikationen. Die internationale bzw. deutschsprachige Bibliothekslandschaft wird von einigen Klassifikationssystemen dominiert, die im folgenden genauer vorgestellt werden sollen. Dabei handelt es sich um die

- Dewey Decimal Classification (DDC),
- Universal Decimal Classification (UDC),
- Library of Congress Classification (LCC),
- Regensburger Verbundklassifikation (RVK),
- und die Basisklassifikation (BK).

Bei den sich anschließenden Ausführungen wird häufiger auf eine an der Universität Wien durchgeführte Studie Bezug genommen, bei der eine Reklassifizierung größerer Bibliotheksbestände der Germanistik das Ziel war. In deren

Verlauf wurden mehrere Klassifikationssysteme untersucht. Die Durchführung der Studie lag in den Händen von Dr. Otto Oberhauser, Oberrat im Österreichischen Bundesministerium für Bildung, Wissenschaft und Kultur, und Wolfram Seidler, Leiter der Fachbibliothek für Germanistik an der Universität Wien.

### **1.3.1 Die Dewey Decimal Classification (DDC)**

Die Dewey Decimal Classification (DDC) ist das international am weitesten verbreitete Klassifikationssystem. Etwa 200.000 Bibliotheken in mehr als 135 Ländern benutzen die DDC (vgl. Oberhauser / Seidler 2000, 11) und allein in 59 Nationalen Bibliotheken findet die DDC Anwendung, wovon 17 in afrikanischen Staaten, 13 in der amerikanisch-karibischen Region, 7 in Europa, 7 im Mittleren Osten und 15 in der asiatisch pazifischen Region liegen (vgl. Mitchell 2000, 88). Neben Buchaufstellungen und Bibliothekskatalogen werden auch Buchhandelsverzeichnisse nach ihr klassifiziert. Übersetzt wurde die DDC mittlerweile in über 30 Sprachen (vgl. OCLC Forest Press<sup>1</sup>).

Der Name der DDC geht zurück auf Melvil Dewey, der 1876 anonym die Schrift „A Classification and Subject Index for Cataloguing and Arranging the Books and Pamphlets of a Library“ veröffentlichte. Darin präsentierte er eine völlig neue Art der Klassifikation von Büchern in Bibliotheken. Nachdem bis dahin die einzelnen Werke mit der entsprechenden Regalnummer ihres Standortes benannt worden waren, schlug er ein System vor, in dem die Bücher in Relation zueinander und nicht zu ihrem physischen Standort ausgezeichnet werden sollten. Als Grundlage sollte das Dezimalsystem dienen, welches 999 Klassen zur Verfügung stellte (vgl. Rowley / Farrow 2000, 215f).

Die DDC basiert auf Prinzipien, die sich als Werkzeug für die Verwaltung von allgemeinen Wissensbeständen eignen: Sachgebiete repräsentierende Notationen in arabischen Zahlen, klar unterscheidbare Kategorien, ausgereifte Hierarchien und ein reiches Netzwerk an Verbindungen zwischen einzelnen Themen. Die Grundklassen bestehen aus wissenschaftlichen Disziplinen oder Wissensgebieten. Auf der untersten Stufe finden sich in der DDC zehn Hauptklassen, die zusammen das gemeinsame Weltwissen abdecken sollen. Jede der Haupt-

klassen ist in weitere zehn Teilbereiche unterteilt und jeder Teilbereich wiederum in zehn Sektionen. Nicht alle Zahlen sind mit Themengebieten belegt (vgl. OCLC Forest Press<sup>1</sup>).

Die zehn Hauptklassen der DDC sind zur Zeit die folgenden:

- 000 Computers, information, & general reference
- 100 Philosophy & psychology
- 200 Religion
- 300 Social sciences
- 400 Language
- 500 Science
- 600 Technology
- 700 Arts & recreation
- 800 Literature
- 900 History & geography

(Quelle: OCLC Forest Press<sup>2</sup>)

Die zweite Ziffer in jeder dreistelligen Nummer steht für einen Teilbereich. So repräsentiert beispielsweise 500 generell die Naturwissenschaften, 510 die Mathematik, 520 die Astronomie und 530 die Physik. Für alle zehn Hauptgebiete ergeben sich so zusammen 100 Bereiche. Auf der dritten Ziffernebene gibt es schließlich insgesamt 1000 Themengebiete. Unter der 530 für Physik finden sich hier u.a. die 531 für klassische Mechanik, die 532 für Mechanik flüssiger Stoffe und die 533 für Mechanik gasförmiger Stoffe. Nach jeder dreistelligen Zahl folgt in der DDC ein Punkt, an den sich bei Bedarf eine weitere Unterteilung in zehn Klassen anschließen kann (vgl. OCLC Forest Press<sup>1</sup>). Auf diese Weise spiegelt sich in der Länge der Notationen bzw. der Anzahl der mit Ziffern (außer der Null) belegten Stellen in der Notation die Hierarchie wieder. Zahlen gleicher Länge sind kürzeren untergeordnet und längeren übergeordnet.

- 600 Technology (Applied sciences)
  - 630 Agriculture and related technologies
    - 636 Animal husbandry

### 636.7 Dogs

(Quelle: OCLC Forest Press<sup>1</sup>)

Ein Themengebiet kann in mehreren Wissensgebieten auftauchen. So findet sich beispielsweise das Thema „Kleidung“ unter 155.95 als Teilgebiet der Psychology (psychologischer Einfluss von Kleidung), außerdem wird es unter 391 unter „Sitten und Gebräuche“ geführt und im Hinblick auf modisches Design trifft man dieses Themengebiet auch unter 746.92 als Teil des Wissensgebietes Kunst an (vgl. OCLC Forest Press<sup>1</sup>).

Publiziert wird das System DDC von Forest Press, einer Abteilung von OCLC (Online Computer Library Center, Inc.). Eine der großen Stärken der DDC ist, dass sie in einer Nationalbibliothek entwickelt wurde, der Library of Congress (LoC). Dort findet sie, neben einem eigenen Klassifikationschema der LoC (siehe Kapitel 1.3.3), weiterhin Anwendung. Allerdings werden in der LoC fast ausschließlich Titel in westeuropäischen Sprachen nach der DDC erschlossen (vgl. Oberhauser / Seidler 2000, 11), insgesamt sind es aber immerhin mehr als 110.000 Titel, die dort jährlich auf diese Weise klassifiziert werden (vgl. OCLC Forest Press<sup>1</sup>).

Das Dewey Editorial Office, welches sich direkt in der LoC befindet, ist dafür zuständig, neue Trends in der Literaturlandschaft ausfindig zu machen und gegebenenfalls Vorschläge für Veränderungen der DDC zu erarbeiten. Diese werden an das Decimal Classification Editorial Policy Committee (EPC), einem zehnköpfigen Gremium, weitergeleitet; alle sieben bis acht Jahre erscheint eine Neuausgabe, wobei aber jeweils nicht mehr als zwei Hauptklassen umfassend erneuert werden. Ein Großteil der DDC ist aber schon seit längerer Zeit nicht mehr revidiert worden (vgl. Oberhauser / Seidler 2000, 11). Hier offenbart sich laut Oberhauser und Seidler auch eine Schwäche der DDC:

„Aufgrund der mit dem Dezimalsystem prinzipiell verbundenen Restriktionen (Zwang zur Einteilung des Wissens in zehn Hauptklassen mit wiederum jeweils zehn Unterklassen usw.) ist die Grundstruktur zum Teil veraltet und nicht aktualisierbar“ (Oberhauser / Seidler 2000, 11).

Die Dewey Decimal Classification ist in zwei Versionen verfügbar, einer ausführlichen und einer gekürzten. Nach Angaben von Oberhauser und Seidler wird die DDC im deutschsprachigen Raum bislang nicht angewendet, sondern ist traditionell eher im anglo-amerikanischen Raum anzutreffen (vgl. Oberhauser / Seidler 2000, 11).

### **1.3.2 Die Universal Decimal Classification (UDC)**

Die Universal Decimal Classification (UDC) wurde als Erweiterung der DDC 1894 entwickelt. Zwei Belgier, Paul Otlet und Henri La Fontaine, veröffentlichten zwischen 1904 und 1907 ihre erste Ausgabe in französischer Sprache (vgl. Rowley / Farrow 2000, 229). Seitdem wurde die UDC ständig verändert und entwickelte sich zu einem System für die Verwaltung bibliographischer Bestände jeglicher Art. Sie lehnt sich weiterhin stark an die DDC an, verfügt aber zusätzlich über eine Fülle von Möglichkeiten, Wissensgebiete weitaus detaillierter und feiner strukturiert als die DDC darzustellen (vgl. McIlwaine 2000, 100).

Die Notationen der UDC sind denen der DDC sehr ähnlich, werden aber um zusätzliche Trenn- und Sonderzeichen wie Klammern, Gleichheitszeichen, Anführungszeichen, Pluszeichen u.a. erweitert. Dadurch werden die Notationen nach Meinung von Oberhauser und Seidler „mit zunehmender Länge unübersichtlich, schwer verständlich und sind schwierig zu bilden“ (Oberhauser / Seidler 2000, 13). Es existieren ebenfalls zehn Hauptklassen; im Gegensatz zur DDC werden die direkt untergeordneten Sachgebiete aber nicht mit Nullen zu dreistelligen Zahlen erweitert:

- 8 Language. Linguistics. Literature
- 80 General questions. Prosody
- 81 Linguistics and languages
- 82 Literature generally
- 821 Literatures of individual languages

(Quelle: Oberhauser / Seidler 2000, 13)

Große Anwendung findet die UDC im naturwissenschaftlichen und technischen Kontext (vgl. Rowley / Farrow 2000, 232). Außerdem kommt sie in stark spezialisierten Institutionen zum Einsatz. Die UDC wird weltweit verwendet, sie wurde in 23 verschiedenen Sprachen veröffentlicht (vgl. UDC Consortium). Auch im deutschen Sprachraum wurde sie insbesondere im Bereich der Spezialbibliotheken und Dokumentationsstellen häufig eingesetzt. In den letzten Jahren ist hier aber eine stark rückläufige Entwicklung zu verzeichnen. Auch im internationalen Bereich lässt die institutionelle Anwendung nach. Als gravierender Nachteil erweist sich, dass die UDC im Gegensatz zur Library of Congress Classification (LCC) (siehe Kapitel 1.3.3) und der DDC keine laufende Überprüfung durch eine große Institution mit eigenen Buchbeständen erfährt (vgl. Oberhauser / Seidler 2000, 13).

### **1.3.3 Die Library of Congress Classification (LCC)**

Die größte Bibliothek der Welt, die Library of Congress (LoC) in Washington D. C., hat ein nach ihr benanntes, eigenes Klassifikationssystem aufgebaut. Ursprünglich für die speziellen Bedürfnisse der LoC entwickelt, erwies sich die LCC auch zur Verwaltung anderer großer wissenschaftlicher Bestände als geeignet. Schon ab 1901 konnte der gedruckte Katalog von anderen Büchereien erworben werden (vgl. Rowley / Farrow 2000, 225f). Im Laufe eines Jahrhunderts hat sich die LCC als ein äußerst praktisches System herausgestellt und die Anwendung beschränkt sich mittlerweile nicht mehr nur auf die USA; auch in anderen Teilen der Welt, insbesondere den englischsprachigen, wurde die Effektivität der Klassifikation erkannt (vgl. Chan / Hodges 2000, 113). Im deutschsprachigen Raum wird die LCC bisher kaum verwendet, ein Anwendungsbeispiel ist die Fremddatenübernahme bei der Universitätsbibliothek Göttingen (vgl. Oberhauser / Seidler 2000, 9).

Die LCC unterteilt das Wissen in 21 unterschiedliche Kategorien oder Klassen. Jede der Klassen wird durch einen Buchstaben des Alphabets bezeichnet. Untergeordnete Klassen werden durch Hinzufügen eines weiteren Buchstabens gekennzeichnet. Weitere Unterklassen ergeben sich durch das Anhängen von

16

Zahlen. Ein Ausschnitt der Oberklasse „Social Science“ verdeutlicht die Struktur:

H Social Sciences

- H Social Sciences (General)
- HA Statistics
- HB-HC Economics
- HD Industry. Labor
- HE Transportation. Communications
- HF Commerce. Marketing. Advertising
- 5383- Resumes
- 5410- Marketing
- 5500.2- Personnel

(Quelle: British Columbia Institute of Technology Library)

Die Signaturenbildung erfolgt mittels einer sogenannten Cutter-Sanborn-Nummern (CSN), wobei auf Basis standardisierter Tafeln eine Kombination aus dem Anfangsbuchstaben des Familiennamens des Autors und einer Zahl gebildet wird. Dieses Verfahren dient der groben Individualisierung bzw. Alphabetisierung innerhalb einer Systematikstelle (Klasse) (vgl. Oberhauser / Seidler 2000, 9). Allerdings kann dies nicht verhindern, dass sich die LCC nur durch die ständige Anwendung erlernen lässt. Im Gegensatz zur Dewey Decimal Classification fehlen die Grundprinzipien, wodurch sich eine Klassifikation dem Benutzer intuitiv erschließen kann.

Der große Vorteil der LCC liegt hingegen darin, dass Spezialisten die Entstehung jeder Klasse von Anfang an überwacht und deren Aufbau und Hierarchie entwickelt haben, um so die Anforderungen sehr großer Dokumentenbestände erfüllen zu können (vgl. Rowley / Farrow 2000, 226f). Da die LCC in einer wichtigen nationalen Institution beheimatet ist, die relative politische und ökonomische Stabilität garantiert und hohe Qualität erfordert, werden die einzelnen Klassen ständig kontrolliert und bei Bedarf überarbeitet (vgl. Chan / Hodges 2000, 113). Wöchentlich werden Neuerungen in der Zeitschrift „Information Bul-

letin“ bekannt gegeben. Die Revision findet somit kontinuierlicher statt als bei der DDC.

#### **1.3.4 Die Regensburger Verbundklassifikation (RVK)**

Die Regensburger Verbundklassifikation (RVK) wurde ursprünglich als reines Aufstellungssystem der Universitätsbibliothek Regensburg entwickelt. Im Gegensatz zu DDC, UDC und LCC ist sie keine Universalklassifikation im eigentlichen Sinn, sondern eine Sammlung von 34 Fachsystematiken, von denen die Geisteswissenschaften besonders gut ausgebaut sind. Sie werden durch ein einheitliches Notationssystem zu einer Gesamtklassifikation zusammen gefasst. Eine Vielzahl von deutschen Bibliotheken finden sich in dem Verbundsystem wieder, welches über die UB Regensburg koordiniert wird (vgl. Oberhauser / Seidler 2000, 14).

Die einzelnen Fachsystematiken der RVK werden kontinuierlich überarbeitet. Das System funktioniert folgendermaßen: Die RVK hat ca. 140 Anwender, die alle jederzeit Neuvorschläge für Notationen oder Änderungen einbringen können. Somit wird die gesamte RVK laufend optimiert (Informationen einer E-Mail vom 22.04.2002 von Frau Barbara Leiwesmeyer, Leiterin der Teilbibliothek Recht und RVK-Verbundkoordinatorin der Universitätsbibliothek Regensburg).

Die Bildung der Notationen lehnt sich an die LCC an, einleitenden Großbuchstaben folgen drei- bis sechsstelligen Zahlen. Die vollständige Systematik findet sich auch im Internet. Am Beispiel der Informationswissenschaft soll die Struktur ausschnittsweise verdeutlicht werden:



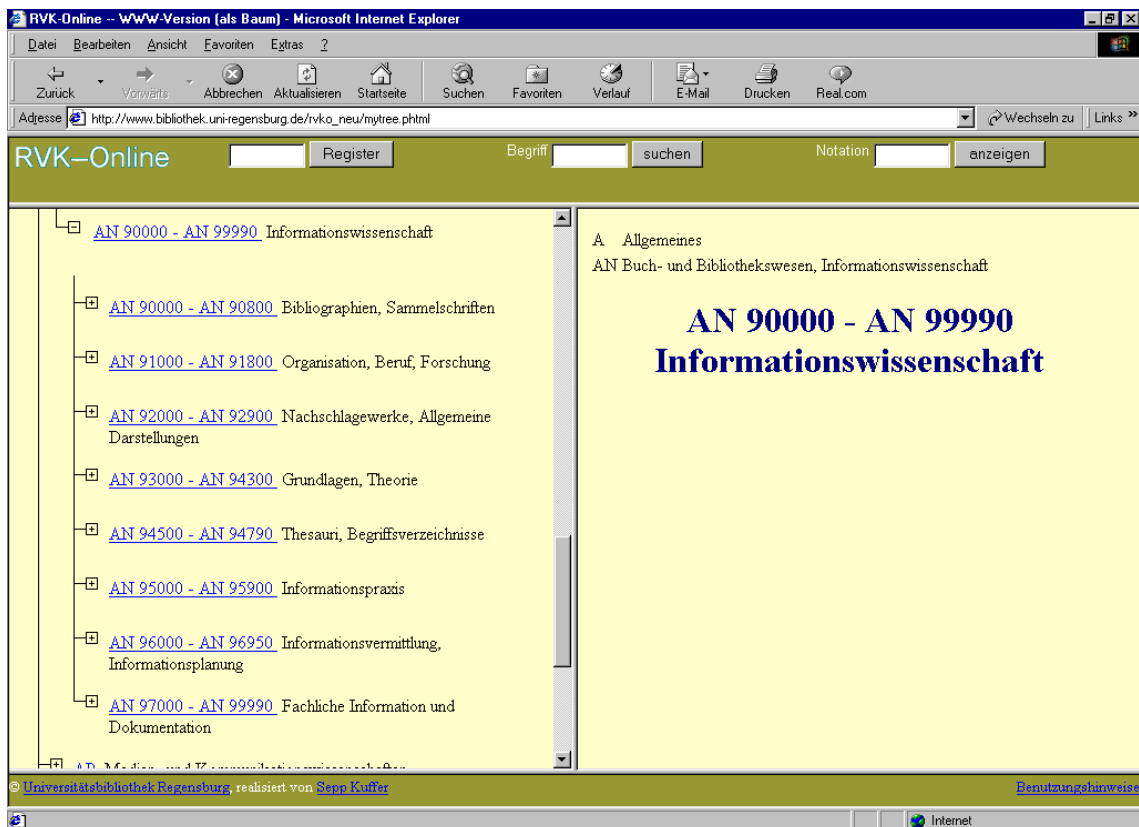


Abbildung 1: Regensburger Verbundklassifikation (RVK), Informationswissenschaft

(Quelle: Universitätsbibliothek Regensburg<sup>1</sup>)

### 1.3.5 Die Basisklassifikation (BK)

Die sogenannte Basisklassifikation (BK) stammt aus den Niederlanden und ist eine Grobsystematik auf aktuellem Wissensstand. Sie wurde primär als Ergänzung zur verbalen Sacherschließung konzipiert und nicht als Aufstellungssystematik.

Eine übersetzte Version des BK wird vom größten Anwender im deutschsprachigen Raum verwendet, dem Gemeinsamen Bibliotheksverbund der Länder Bremen, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen (GBV). Sie kommt im Online-Katalog als Ergänzung der Schlagwortung nach den „Regeln für den Schlagwortkatalog (RSWK)“ zum Einsatz. Auch die Universitätsbibliothek Hildesheim benutzt

zusätzlich neben der eigenen lokalen Klassifikation die BK. Die Pflege und Aktualisierung des BK erfolgt gemeinsam durch niederländische und deutsche Gremien.

Die Basisklassifikation besteht aus 89 Oberkategorien, denen jeweils die entsprechenden Unterkategorien zugeordnet sind. Die Notationen der BK sind vierstellig, wobei die ersten beiden Ziffern (die die Oberkategorien repräsentieren) von den letzten beiden (die für die Unterkategorien stehen) durch einen Punkt getrennt werden. Die Kategorie Informatik zeigt den angesprochenen Aufbau der Systematik:

### 54 Informatik

#### 54.00 Informatik: Allgemeines

Syn.: Computerwissenschaft; Datenverarbeitung

#### 54.01 Geschichte der Informatik

#### 54.04 Ausbildung, Beruf, Organisationen

Verw.: Computer im Unterricht -> 81.68  
(Computereinsatz in Unterricht und Ausbildung)

#### 54.08 Informatik in Beziehung zu Mensch und Gesellschaft

#### 54.10 Theoretische Informatik

(Quelle: Gemeinsamer Bibliotheksverbund)

## **2 Die Problematik der informationswissenschaftlichen Literaturrecherche in der Universitätsbibliothek Hildesheim**

Wie schon mehrfach erwähnt, ist die Situation für Studierende in Hildesheim bei der Suche nach Büchern und anderen Dokumenten aus dem Bereich der Informationswissenschaft unbefriedigend. Die relevanten Bibliotheksbestände sind nicht unter einer Ordnungshierarchie zusammengefasst, sondern stehen verteilt über die verschiedenen Notationen der lokalen Aufstellungssystematik. Sie sind den unterschiedlichsten Disziplinen zugeordnet, in erster Linie der Computerwissenschaft, aber auch der Mathematik, der Bibliothekswissenschaft, der Betriebswissenschaft, der Linguistik, der Rechtswissenschaft u.a. In fast jeder Disziplin finden sich in unterschiedlich großer Anzahl Bücher, die im informationswissenschaftlichen Kontext angesiedelt werden können und deren Repräsentation dort wünschenswert wäre. Die Ursache dafür liegt in der noch sehr jungen Entstehungsgeschichte und der interdisziplinären Ausrichtung der Informationswissenschaft, die somit bei der Aufstellung der Bibliotheken keine Berücksichtigung fand.

Betroffen von dieser Problematik sind in der Hildesheimer Universität insbesondere die Studierenden aus dem Magisterstudiengang Internationales Informationsmanagement (IIM) mit dem Schwerpunkt Informationswissenschaft und dem Bachelor-Studiengang Informationsmanagement / Informationstechnologie (IMIT), deren Studiengänge sich nicht in der Ordnungssystematik der Bibliothek wieder finden. Diesen Studierenden wird die Recherche deutlich erschwert.

Zum einen werden ihnen die Vorteile, die eine lokale Aufstellungssystematik bieten kann, verwehrt. Es gibt keinen Standort, an dem man alle Bücher aus dem Themenbereich Informationswissenschaft überschauen kann. Ein direkter Zugriff, der dem Suchenden einen schnellen Überblick über den Umfang des Buches, seine Gestaltung und seinen Inhalt erlaubt und ihm so eine erste Relevanzentscheidung ermöglicht, ist nicht gegeben.

Zum anderen müssen die Studierenden der Informationswissenschaft, wie auch ihre Kommilitonen aus anderen Studiengängen, mit den Nachteilen von physischen Bibliotheksregalen leben. Aufgrund der Starrheit des Systems ist keine Anpassung auf neue Sichtweisen des Bestandes möglich, sondern der Fächerkanon bei Gründung der Bibliothek wird beibehalten. Auf aktuelle Entwicklungen wird kaum oder nur langsam reagiert. Würde eine neue Kategorie wie die Informationswissenschaft eingefügt, so könnten bereits unter anderen Themenbereichen eingeordnete Bücher kaum berücksichtigt werden, obwohl sie inhaltlich eventuell besser oder zumindest genauso gut hierher passen würden. Einen gravierenden Nachteil für den Benutzer stellen auch die fehlenden Hinweise auf Titel aus nicht frei zugänglichen Beständen, wie z.B. dem Magazin, dar. Gleiches gilt für ausgeliehene Bücher und solche, die für eine gewisse Zeit in Semesterapparaten stehen.

Durch eine fehlende Ordnungssystematik für die Informationswissenschaft entgeht den Studierenden auch die Möglichkeit des Browsens durch den Bücherbestand. Für jede Notation (z.B. CSC 450 oder LIN 310) existiert nämlich in der UB Hildesheim eine Liste von Titeln, die diesem Themenbereich zugeordnet sind. Dies wird dadurch ermöglicht, dass vielen Titeln neben der Signatur sogenannte Nebenstellen zugeordnet sind (unter „Lokale Sachgruppen“), also Notationen, die sich von der der Signatur unterscheiden. In diesen Notationslisten können Titel auch mehrfach aufgeführt werden, da sie nicht mit den lokalen Aufstellungen identisch sind. Auf diese Weise sollen die oben genannten Nachteile lokaler Ordnungssystematiken ausgeglichen werden. Diese Form eines virtuellen Zugriffs auf Literatur soll im Rahmen dieser Arbeit für den Bereich der Informationswissenschaft entwickelt werden.

## **2.1 Verdeutlichung der Problematik anhand von Literaturbeispielen**

Um die auftretenden Probleme bei Suche nach informationswissenschaftlicher Literatur in der Universitätsbibliothek einmal zu verdeutlichen, wurden konkrete Recherchen nach relevanten Titeln durchgeführt. Es sollte belegt werden, dass diese Titel quer über fast alle Signaturen verteilt zu finden sind. Ausgangspunkt

für die Suche war eine bereits erstellte Liste, in welcher informationswissenschaftlichen Begriffen die Notationen der Universitätsbibliothek Hildesheim zugeordnet worden waren, die die jeweiligen Begriffe in der Notationsbeschreibung enthalten (siehe Kapitel 4.1). Aus der Vielzahl der ermittelten Notationen wurden beispielhaft einige ausgewählt und als Suchanfrage in den Web-OPAC der UB Hildesheim eingegeben. Aus der Ergebnisliste wurden wiederum exemplarisch Titel ausgesucht, bei denen der thematische Bezug zur Informationswissenschaft klar erkennbar ist. Insgesamt wurden elf Systemstellen herangezogen, darunter die, die am häufigsten in der Liste auftauchen (CSC, BUB, BWL, LIN, MAT, TFF, ERZ, REC). Die Titel mit den Signaturen wurden in einer Tabelle zusammengetragen:

<b>Titel</b>	<b>Autor</b>	<b>Signatur</b>
Datenstrukturen und Algorithmen	Ralf Hartmut Güting	CSC 410 : G74
From Gutenberg to the global information infrastructure : access to information in the networked world	Christine L. Borgman	BUB 183 : B56
Die Zukunft des Online-Marketing : eine explorative Studie über zukünftige Marktkommunikation im Internet	Volker Walter	BWL 573 : W05
Language and computers : a practical introduction to the computer analysis of language	Geoff Barnbrook	LIN 410 : B15
Bundesdatenschutzgesetz : [mit Erläuterungen] ; BDSG	Hans-Joachim Ordemann; Rudolf Schomerus, Peter Gola	REC 480 : O64,b
Virtuelle Universität - virtuelles Lernen	Rolf Schulmeister	ERZ 857 : S11
Computereinsatz im Geographieunterricht	Bruno Pohl	GEO 671 : P55
Internet für Historiker : eine praxisorientierte Einführung	Andreas Ohrmund; Paul Tiedemann	GES 084 : O33
Fuzzy multiple objective decision making : methods and applications	Young-Jou Lai; Ching-Lai Hwang	MAT 124 : L03
Die Mondlandung des Internet : die Bundestagswahl 1998 in den elektronischen Kommunikationsforen	Rainer Kuhlen	POL 351 : K74
Multimedia-Kommunikation : Theorien, Trends und Praxis	Peter Ludes	TFF 140 : M75

Tabelle 1: Recherche nach Hildesheimer Titeln über verschiedene Notationen

Hinter den Kürzeln der Signaturen verbergen sich die folgenden Themengebiete der Systematiken:

- CSC: Informatik (computer science)
- BUB: Buch & Bibliothek, Information & Dokumentation
- BWL: Wirtschaftswissenschaften
- LIN: Linguistik
- REC: Rechtswissenschaft
- ERZ: Erziehungswissenschaft
- GEO: Geographie, Geowissenschaften, Bergbau
- GES: Geschichte
- MAT: Mathematik
- POL: Politikwissenschaft, Verwaltung
- TFF: Medienwissenschaften & Publizistik

Beim Betrachten der Ergebnistabelle fällt auf, dass abgesehen von dem Buch von Christine L. Borgmann (deren Buch aber durch die Thematik digitaler Bibliotheken inhaltlich dorthin passt) alle Titel sowohl einen deutlichen Bezug zur Informationswissenschaft als auch zu dem jeweiligen Schwerpunktthema ihrer Signatur aufweisen. Somit ist die entsprechende Einordnung und Aufstellung voll gerechtfertigt. Das Auffinden dieser Bücher ist allerdings nur dann möglich, wenn der Suchende ganz gezielt danach forscht oder zufällig darauf trifft, bzw. wenn ihm bekannt ist, an welchem Ort in der betreffenden Aufstellungsnotation (z.B. GES 084) eventuell informationswissenschaftliche Bücher zu finden sind. Einen Überblick kann er sich auf diese Weise jedenfalls nicht verschaffen.

Eine weitere Erkenntnis dieser Untersuchung ist, dass Schlagwörter bei der Suche zwar helfen, dass die Thesauri aber auch Lücken aufweisen, die relevante Titel vom Suchergebnis ausschließen können. Beispielsweise wurde das Buch von Volker Walter („Die Zukunft des Online-Marketing: eine explorative Studie über zukünftige Marktkommunikation im Internet“) mit den Begriffen Telemarketing, Internet und Prognose verschlagwortet. Der allgemein gebräuchliche Term „Online-Marketing“ fehlt, müsste aber bei einer Suche als Eingabe erwartet werden (siehe Abbildung 2). Zwar würde eine Suche mit diesem Term

## 2 Die Problematik der informationswissenschaftlichen Literaturrecherche in der Universitätsbibliothek Hildesheim

über die Titelstichwörter in diesem konkreten Fall erfolgreich sein, bei einem anderen Fall ohne das entsprechende Suchwort in der Titelzeile wäre das Suchergebnis negativ.

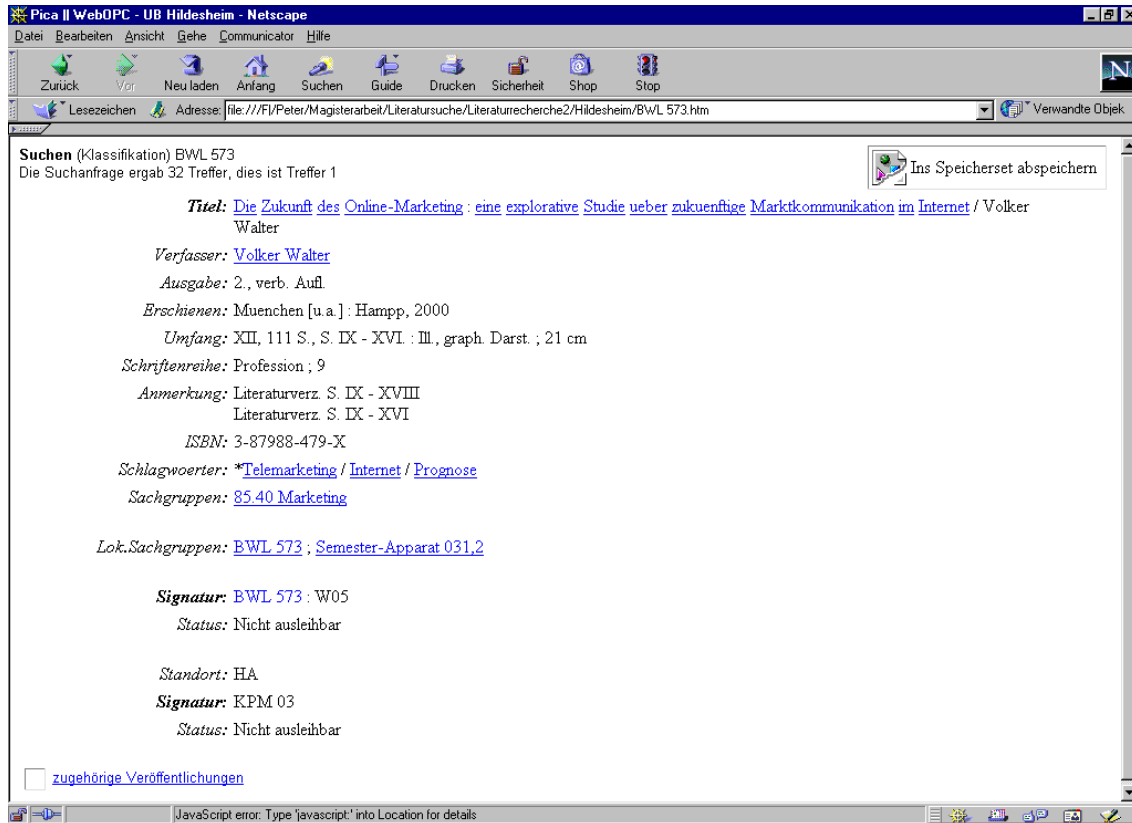


Abbildung 2: Trefferanzeige Web-OPAC UB Hildesheim, Volker Walter  
(Quelle: Universitätsbibliothek Hildesheim<sup>1</sup>)

Ein weiteres Beispiel für diese Problematik ist das Buch „Die Mondlandung des Internet: die Bundestagswahl 1998 in den elektronischen Kommunikationsforen“ von Rainer Kuhlen. Obwohl schon im Titel das Wort „Internet“ auftaucht und sich das Buch auch ganz ausführlich mit den Internet-Auftritten der Parteien während des Bundestagswahlkampfes 1998 beschäftigt, fehlt der Begriff bei den zugeordneten Schlagwörtern (siehe Abbildung 3). Mit den Schlagwort-Suchtermen „Deutschland“, „Elektronisches Forum“, „Bundestagswahl“, „Wahlkampf“ und „Geschichte 1998“ würde man das Buch von Kuhlen finden, nicht aber mit „Internet“. Für die Suche mit Titelstichwörtern gilt das oben Gesagte.

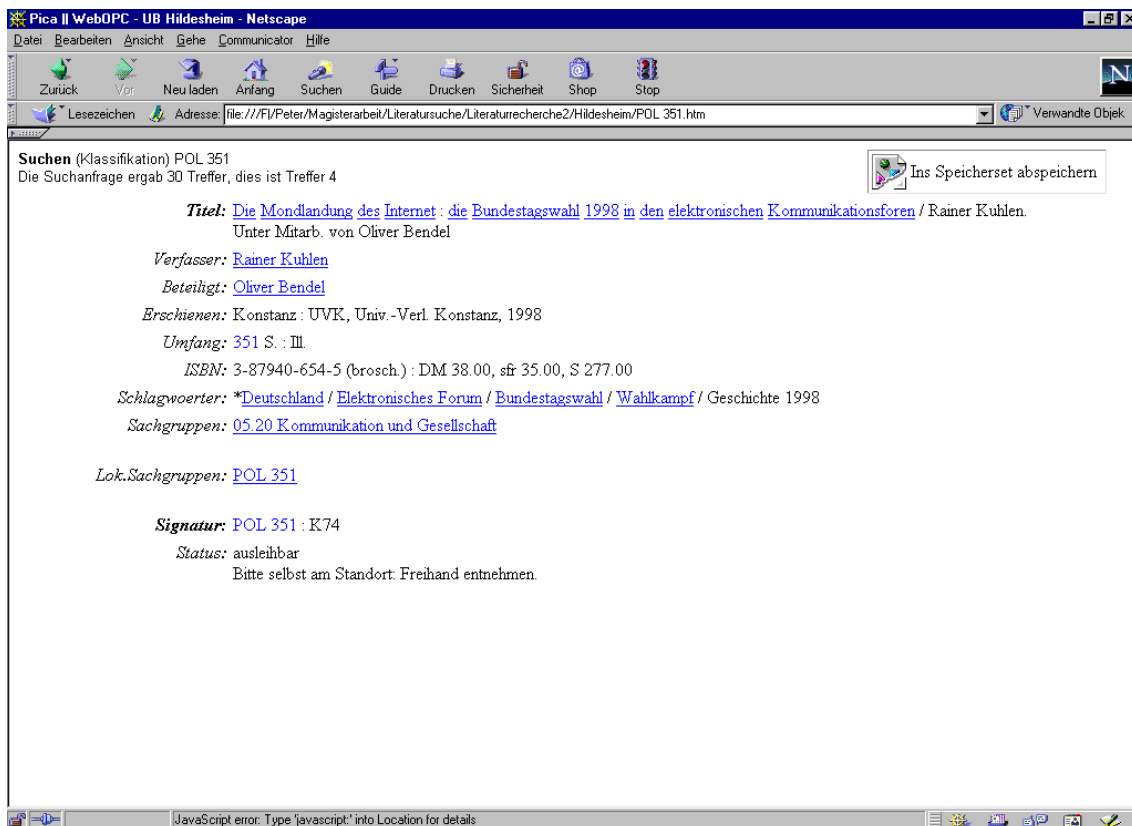


Abbildung 3: Trefferanzeige Web-OPAC UB Hildesheim, Rainer Kuhlen  
(Quelle: Universitätsbibliothek Hildesheim<sup>1</sup>)

Weiter ist zu beobachten, dass in nur geringem Ausmaß unter „Lokale Sachgruppen“ Nebenstellen mit der Notation CSC eingetragen sind, unter der man noch am ehesten eine Aufstellung informationswissenschaftlicher Bücher erwarten würde. Somit hätte auch dort ein Benutzer schlechte Chancen, durch einfaches Browsen auf das gesuchte Dokument zu treffen. Der Titel „Multimedia-Kommunikation: Theorien, Trends und Praxis“ von Peter Ludes wird beispielsweise nur unter den lokalen Sachgruppen TFF 140, TFF 780 und TFF 789 geführt (siehe Abbildung 4), obwohl sich unter CSC 650-659 Titel zum Themengebiet „Multimedia“ finden (vgl. Universitätsbibliothek Hildesheim<sup>2</sup>).



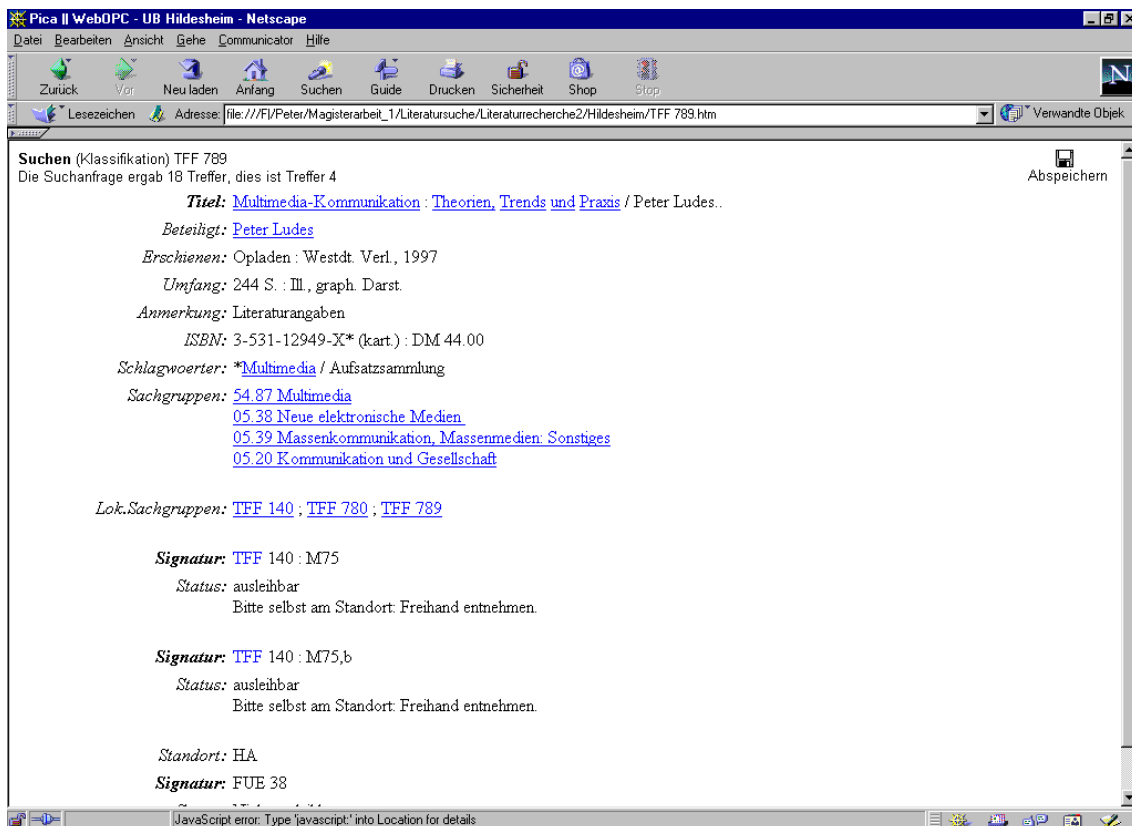


Abbildung 4: Trefferanzeige Web-OPAC UB Hildesheim, Peter Ludes  
(Quelle: Universitätsbibliothek Hildesheim<sup>1</sup>)

## 2.2 Vergleiche mit anderen Bibliotheken mit eigener informationswissenschaftlicher Klassifikation

### 2.2.1 Auswahl von geeigneten Bibliotheken bzw. Klassifikationen

Für einen Vergleich der Situation in der Hildesheimer Universitätsbibliothek mit der in anderen Bibliotheken, deren Hochschulen über einen informationswissenschaftlichen Studiengang verfügen, sollten zuerst einmal relevante Standorte ermittelt werden, die über eine eigene Klassifikation informationswissenschaftlicher Themen- und Lehrgebiete verfügen. Außerdem sollten, im Hinblick auf die spätere Erstellung einer eigenen virtuellen Systematik für die UB Hil-

desheim, Institutionen und Organisationen mit informationswissenschaftlichem Bezug in Hinsicht auf eine Katalogisierung der Inhalte untersucht werden.

Eingehend geprüft wurden Standorte informationswissenschaftlicher Studiengänge im deutschsprachigen Raum, der Hochschulverband Informationswissenschaft (HI), die Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis (DGI), die Gesellschaft für Informatik e.V. (GI), die American Society for Information Science & Technology (ASIST), die Association for Computing Machinery (ACM) und das Internationale Begegnungs- und Forschungszentrum für Informatik in Schloss Dagstuhl.

Drei der untersuchten Institutionen und Verbände weisen umfangreichere Klassifikationen informationswissenschaftlicher Dokumente auf:

- Die Konstanzer Universitätsbibliothek stellt eine eigene systematische Gliederung zur Verfügung, das „kid“ (Kybernetik, Informatik, Datenverarbeitung und Informationswissenschaft) (siehe beiliegende CD-ROM). An der Konstanzer Universität wird der Diplomstudiengang Informationswissenschaft angeboten.
- Die Universitätsbibliothek Regensburg verwaltet in der Regensburger Verbundklassifikation (RVK) (siehe beiliegende CD-ROM) einen Themenbereich „Informationswissenschaft“. An der zugehörigen Hochschule existiert das Fach als Teilbereich des MA-Studienganges „Allgemeine Sprachwissenschaft“.
- Die Universität des Saarlandes in Saarbrücken bietet eine Institutsbibliothek mit eigener Klassifikation für die Fachrichtung „Informationswissenschaften“ an (siehe beiliegende CD-ROM). Studierende können diese im Magisterstudium als Haupt- oder Nebenfach, oder als Ergänzungsfach bei bestimmten Diplomstudiengängen belegen.

Die Universitäts- und Landesbibliothek (ULB) Düsseldorf, deren Hochschule den Studiengang „Informationswissenschaft“ als MA-Nebenfach anbietet, stellt im Rahmen ihrer Fachsystematik (vgl. Universitäts- und Landesbibliothek Düsseldorf<sup>1</sup>) im Bereich „Buch-, Bibliotheks- und Informationswesen“ eine Rubrik

„Informationswissenschaft und Dokumentation“ zur Verfügung. Dort findet sich eine Auflistung von 39 einzelnen Themengebieten, denen direkt die entsprechende Literatur zugeordnet ist. Es fehlen also, anders als bei den oben genannten Klassifikationen, weitere Unter- und Teilrubriken. Außerdem ist festzustellen, dass sich unter den 39 Themengebieten allein 21 befinden, die man unter „Allgemeines“ und „Grundlagen/Theorie“ fassen kann. Weitere zehn Themengebiete sind einzelne wissenschaftliche Disziplinen, in denen Dokumentationsverfahren zur Anwendung kommen („Geisteswissenschaften“, „Medizin“, „Technik“, usw.). Die verbleibenden acht Wissensgebiete dieser Aufzählung finden sich als Unter- oder Oberrubriken in den Klassifikationen der Standorte Konstanz, Regensburg und Saarbrücken wieder; dort werden sie allerdings, im Gegensatz zur Düsseldorfer Klassifikation, in einen inhaltlichen Zusammenhang gestellt.

Auch die Gesamtzahl der in diesem Abschnitt der Düsseldorfer Fachsystematik geführten Bücher ist mit 1.274 Titeln (Stand: 29.01.2002) eher gering (Der Umfang an informationswissenschaftlicher Literatur in Konstanz, Regensburg und Saarbrücken wird an späterer Stelle untersucht.). Auf den Seiten der ULB Düsseldorf heißt es zu dieser Problematik:

„Literatur zum Fach „Informationswissenschaft“ ist in der Universitäts- und Landesbibliothek Düsseldorf im Freihandbereich vor allem im Bereich bib (d.h. Buch, Information und Bibliothek) aufgestellt. Die Inhalte des Faches sprengen aber diesen mehr traditionell bibliothekswissenschaftlichen und –geschichtlichen Bestand und sind häufig auf Systemstellen in Psychologie, Soziologie, Datenverarbeitung und Wirtschaftswissenschaften zu finden“ (Universitäts- und Landesbibliothek Düsseldorf<sup>2</sup>).

Um den Studierenden trotzdem zu einer Übersicht über die verfügbare Literatur zu verhelfen, hat man im Rahmen der „Düsseldorfer Virtuellen Bibliothek“ ein alphabetisches Register zu den Beständen des Studienganges „Informationswissenschaft“ erarbeitet. Hier werden in einer alphabetischen Liste informationswissenschaftlichen Themenbereichen die entsprechenden Notationen zu-

geordnet, unter denen man zugehörige literarische Bestände in der ULB Düsseldorf finden kann (vgl. Universitäts- und Landesbibliothek Düsseldorf<sup>2</sup>).

Der Fachbereich Informations- und Wissensmanagement an der FH Darmstadt stellt auf seinen Seiten eine sogenannte „WebSite Methodik“ zur Verfügung (vgl. FH Darmstadt<sup>1</sup>). Auf diesem „Informationsserver“ werden 567 Dokumente aus insgesamt 101 Themenbereichen verwaltet, die nach Themen bzw. nach Form (Lehrveranstaltung, Overhead-Folie, Hauptstudium, usw.) unterteilt sind. An der Art der vorliegenden Dokumente sieht man schon, dass es sich hierbei um im Rahmen des Studiums erstellte Materialien handelt. Außerdem findet man eine zur WebSite Methodik zugehörige thematische Karte (vgl. FH Darmstadt<sup>2</sup>), die eine Klassifikation der verwalteten Dokumente darstellt. Allerdings ist sie in Umfang und Ausführlichkeit nicht mit den Klassifikationen der Standorte Konstanz, Regensburg und Saarbrücken zu vergleichen.

Eine Art Klassifikation der informationswissenschaftlichen Lehrveranstaltungen an den Hochschulen Graz, Darmstadt, Konstanz, Regensburg und Hildesheim hat der Hochschulverband Informationswissenschaft (HI) erstellt (siehe beiliegende CD-ROM). Dort werden in 20 Haupt- und 6 Unterkategorien Vorlesungen, Seminare und Übungen informationswissenschaftlichen Inhalts erfasst. Eine vom HI anvisierte „Virtuelle Akademie Informationswissenschaften“, die ebenfalls eine Systematisierung der Inhalte erfordern würde, wurde bisher offensichtlich noch nicht realisiert.

Die Gesellschaft für Informatik e.V. besitzt eine Aufteilung ihrer Aufgabengebiete nach Fachbereichen (vgl. Gesellschaft für Informatik). Allerdings weisen diese, wie zu erwarten, eine stark informatische Ausrichtung auf und Themen der Informationswissenschaft kommen darin zu kurz. Gleiches gilt für das Computing Classification System der ACM (vgl. ACM).

Die weiteren, in die Untersuchung einbezogenen Organisationen, die DGI, die ASIST und das Begegnungs- und Forschungszentrum für Informatik in Schloss Dagstuhl haben keine eigenen Klassifikationen der angebotenen Inhalte vorzuweisen. Gleiches gilt für den Literaturbestand im Sachgebiet „Informationswis-

senschaft“ an anderen deutschsprachigen Hochschulen, die einen gleich oder ähnlich lautenden Studiengang anbieten.

Nach der durchgeführten Analyse lassen sich folgende Schlüsse ziehen:

- Die informationswissenschaftlichen Klassifikationen der Universitätsbibliotheken in Konstanz, Regensburg und Saarbrücken erscheinen aufgrund ihres Inhaltes, ihres Umfangs und ihrer Detailliertheit als geeignet für weitere Vergleiche und Untersuchungen.
- Sowohl die einzelnen Wissensgebiete der ULB Düsseldorf und des Fachbereiches Informations- und Wissensmanagement an der FH Darmstadt, als auch der Systematik des HI von informationswissenschaftlichen Lehrveranstaltungen können in eine zu erstellende Klassifikation für die Universitätsbibliothek Hildesheim einfließen.
- Das alphabetische Register der ULB Düsseldorf kann als Vorbild für eine eigene ähnliche Liste dienen (siehe Kapitel 4.1).

### ***2.2.2 Vergleich mit der Situation Studierender in Konstanz, Regensburg und Saarbrücken***

In dem nun folgenden Schritt wurde untersucht, ob die Titel aus der in Kapitel 2.1 durchgeführten Literaturrecherche, die per Zufallsprinzip aus den einzelnen Notationen ausgewählt wurden, an den genannten Standorten (Regensburg, Konstanz, Saarbrücken) ebenfalls zu finden sind. Auch die Signatur und eingetragene Nebenstellen sollten untersucht werden. Zu diesem Zweck wurden die jeweiligen OPACs der einzelnen Standorte („Koala“ in Konstanz, Regensburger OPAC und „ViBI“ in Saarbrücken) mit den entsprechenden Suchanfragen kontaktiert.

Dabei stellte sich heraus, dass in Konstanz neun der elf Hildesheimer Titel stehen, in Regensburg sieben und in Saarbrücken nur zwei (siehe Tabelle 2). Außerdem fiel auf, dass von den Konstanzer Titeln nur drei mit einer Signatur des kid versehen waren. Bei den Regensburger Titeln trug kein einziges der sieben Bücher eine Signatur aus der informationswissenschaftlichen Klassifikation (AN = „Buch- und Bibliothekswesen, Informationswissenschaft“). In Saarbrücken

waren beide gefundenen Bücher durch die ersten drei Ziffern in der Signatur als informationswissenschaftliche Titel gekennzeichnet.

<b>Autor</b>	<b>Konstanzer Signatur</b>	<b>Regensburger Signatur</b>	<b>Saarbrücker Signatur</b>
Ralf Hartmut Güting	kid 112/ g92 wrp 328.02/ w16		
Christine L. Borgman	kid 818/ b67		
Volker Walter	wrp 328.02/ w16	40/QP 650 W234(2)	
Geoff Barnbrook			
Hans-Joachim Ordemann; Rudolf Schomerus, Peter Gola	juf 550:r/ o72(6)	31/PZ 4300 O65(6)	
Rolf Schulmeister	erz 793/ s24k	72/ST 420 S386 V8	
Bruno Pohl		00/RB 10865 P748	
Andreas Ohrmund; Paul Tiedemann	gsb 995/ o47	00/NB 2800 O38	
Young-Jou Lai; Ching-Lai Hwang	sta 630:i/l22- 404	40/QH 233 L185	
Rainer Kuhlen	soz 798/ k94 kid 228.50/ k89		030-4324
Peter Ludes	sf 3026	406/ST 280 L944	250-4270

Tabelle 2: Erfassung der Hildesheimer Titel in Klassifikationen der anderen Standorte

Daraus kann man die vorsichtige Vermutung ableiten, dass die Literaturversorgung im Bereich Informationswissenschaften in Konstanz und Regensburg ähnlich gut zu sein scheint, wohingegen der Bestand in Saarbrücken wohl deutlich geringer ist. Außerdem muss man feststellen, dass selbst in Universitätsbibliotheken mit einer gesonderten Erfassung und Ausweisung der informationswissenschaftlichen Titel nicht alle relevanten Exemplare Eingang in die für sie bestimmte Klassifikation finden.

Die Situation für Studierende der Informationswissenschaft an den untersuchten Standorten ist also insofern besser als für die Studierenden in Hildesheim, als

dort eine Klassifizierung der Literatur vorliegt. Allerdings wäre es von großem Nachteil, sollte die informationswissenschaftliche Literatur in derart unzureichender Weise erfasst sein, wie es das Ergebnis der Recherche vermittelt. Durch die Existenz einer Klassifikation erhält ein Benutzer schließlich den Eindruck, er wäre im Besitz eines umfassenden Überblickes über die zur Verfügung stehende Literatur in den einzelnen Teilbereichen der Informationswissenschaft.

Aus diesem Grund war es notwendig, im Rahmen einer Evaluierung an diese Recherche mit wenigen Titeln eine sehr umfangreiche Literaturrecherche anzuschließen, um so ein genaueres Bild davon zu bekommen, wie sich die Situation in Konstanz, Regensburg und Saarbrücken darstellt.

### **3 Evaluierung ausgewählter Bibliotheken mit dem Sachgebiet Informationswissenschaft**

In den Kapiteln 2.2.1 und 2.2.2 wurden bereits einige Vermutungen und Einschätzungen in Bezug auf die klassifikatorische Erfassung und Ausstattung an den anderen Standorten der Informationswissenschaft in Konstanz, Regensburg und Saarbrücken geäußert. Um, vor dem Hintergrund eines späteren Abgleiches des Hildesheimer Bestandes und der Erstellung einer eigenen Klassifikation, zu einer genaueren und aussagekräftigeren Beurteilung zu kommen, wurde eine umfangreiche Recherche mit informationswissenschaftlicher Literatur durchgeführt. Durch diese stichpunktartige Untersuchung sollte u.a. ermittelt werden,

- wie gut die Gesamtausstattung der einzelnen Bibliotheken ist,
- wie umfangreich einzelne Themengebiete ausgestattet sind,
- wie aktuell die Titel der Themengebiete sind,
- wieviele der gefundenen Titel in den Klassifikationen erfasst sind,
- an welcher Stelle die Bücher in den Klassifikationen zu finden sind
- und welche Kongruenzen zwischen den Klassifikationen bestehen.

Am Ende sollte eine Aussage darüber getroffen werden, welche Klassifikationen am besten dazu geeignet sind, einen Überblick über den informationswissenschaftlichen Bücherbestand zu geben und somit einen schnellen Zugriff für den Suchenden zu ermöglichen.

Zu diesem Zweck wurden für die Standorte Hildesheim, Konstanz, Regensburg und Saarbrücken jeweils fünf unterschiedliche Themenbereiche der Informationswissenschaft festgelegt. Aus diesen fünf Themengebieten wurden insgesamt 50 Titel ausgesucht, deren Zusammenstellung zufällig war und ohne eine Vorauswahl erfolgte. Entscheidend war lediglich die Zuordnung zu einem bestimmten Teilbereich der Informationswissenschaft; die Auswahl der Themengebiete orientierte sich an den Rubriken der jeweiligen Klassifikation des Standortes. Eine gewisse Breite der Themen sollte dabei gewährleistet werden. Am Ende kamen somit 200 Titel zusammen; mit jeweils 150 wurde in jedem



Web-OPAC der vier Bibliotheken geforscht. Die Aufteilung hatte folgendes Aussehen:

- Hildesheim
  - E-Commerce
  - Informationsmanagement
  - Information Retrieval
  - Maschinelle Übersetzung / Computerlinguistik
  - Mensch-Maschine-Interaktion
- Konstanz
  - Informationspsychologie
  - Informationsrecht
  - Informationstechnologie (Software)
  - Informationsverhalten
  - Klassifizieren
- Regensburg
  - Informationsnutzung
  - Informationsspeicherung
  - Informationssysteme
  - Informationsvermittlung
  - Klassifikationssysteme
- Saarbrücken
  - Datenbanksoftware
  - Elektronisches Publizieren
  - Evaluation von Informationssystemen
  - Expertensysteme
  - Informationsnetze

Die Titel und Signaturen der gesuchten und gefundenen Bücher wurden nach Durchführung der Recherche in einer Tabelle eingetragen (Auszug: siehe Anlage 1, komplette Tabelle: siehe beiliegende CD-ROM). Sowohl auf die einzelnen

Themenbereiche als auch auf die Standorte bezogen erfolgte eine mengenmäßige Erfassung.

### **3.1 Auswertung der Literaturrecherche**

#### **3.1.1 Gesamtausstattung**

Als erste Erkenntnis der Recherche lässt sich festhalten, dass von den vier verglichenen Standorten die Universitätsbibliothek Konstanz den umfangreichsten Buchbestand im Bereich informationswissenschaftlicher Titel aufweist. Von den 200 gesuchten Titeln fanden sich 133 in der UB Konstanz wieder (siehe beiliegende CD-ROM), wobei man natürlich berücksichtigen muss, dass 50 von diesen Titeln aus dem kid für die Recherche ausgewählt worden waren. Von 150 Büchern der anderen Standorte wurden somit 83 in Konstanz gefunden, was einer prozentualen Trefferquote von 56 % entspricht. Legt man als Basis für die Berechnung die Gesamtmenge der 200 Titel zugrunde, so ergibt sich ein Prozentsatz von 67. Herr Gerhard Schmitz-Veltin, in der Universitätsbibliothek Konstanz Fachreferent für Physik und Mathematik und zuständig für das kid, bestätigte in einer E-Mail vom 12. Dezember 2001, dass das kid mehr als 23.000 Titel enthält. Hinzu kommen ca. 2000 Medien und etwa 10.000 Bände Zeitschriften. Eine zugesandte Datei mit den Signaturen und Kurztiteln des kid enthielt 23.697 Einträge (Print- und Digitalmedien). Daneben existiert eine Menge unbekannter Größe an Titeln, die die Signatur eines anderen Sachgebietes besitzen, als Systemstelle (entspricht einer Nebenstelle in Hildesheim) aber eine Notation des kid eingetragen haben.

Einen etwas geringeren Bestand weist die Universitätsbibliothek der Stadt Regensburg auf. Hier konnten von 150 informationswissenschaftlichen Büchern fremder Standorte 60 erfolgreich recherchiert werden, was einem Prozentsatz von 40 entspricht (Bezogen auf 200 Titel sind es 55 Prozent.). Allerdings liegt hier das Problem in einem geringen Erfassungsgrad informationswissenschaftlicher Bücher in der dafür vorgesehenen Klassifikation. Der Gesamtbestand an

Büchern, die dort unter den Notationen AN 90000 – AN 99000 abgelegt sind, betrug am 19.04.2002 nur 970 Titel (vgl. Universitätsbibliothek Regensburg<sup>2</sup>).

An dritter Stelle bei der mengenmäßigen Auswertung des Recherche-Ergebnisses steht die UB Hildesheim. 30 Bücher von 150 Suchanfragen mit Titeln anderer Standorte bzw. 80 von 200 Titeln wurden gefunden (20% bzw. 40 %). Am schlechtesten schnitt die Institutsbibliothek Informationswissenschaft der Universität Saarbrücken ab, wo sich nur 18 von 150 bzw. 68 von 200 Titeln im Bestand wiederfanden (12% bzw. 34%). Im „ViBI“ (Virtuelles Bibliotheksinformationssystem) können Benutzer auf 2.524 Bücher und 121 Zeitschriften zurückgreifen.

### 3.1.2 Ausstattung der einzelnen Themengebiete

Um die Ausstattung der Bibliotheken stichpunktartig in einzelnen Teilbereichen zu untersuchen, wurden die Treffermengen der Literaturrecherche in einer Tabelle zusammengetragen. Die Zahl in Klammern hinter dem Themengebiet gibt die Gesamtmenge an Titeln wieder, die bei der Suchanfrage mit der Notation eines Themengebiet in der betreffenden Bibliothek erzielt wurde; aus dieser wurden dann die insgesamt 50 Titel von jedem der vier Standorte ausgewählt, die als Grundlage für die Recherche in den anderen Büchereien dienten (Da diese keine Suchtreffer im eigentlichen Sinn darstellen, sind die entsprechenden Zahlen kursiv gedruckt). Die Hildesheimer Zahlen ergeben sich aus kombinierten Titelwörter- und Schlagwort-Suchabfragen an den WEB-OPAC der Universitätsbibliothek.

	<b>Hildesheim</b>	<b>Konstanz</b>	<b>Regensburg</b>	<b>Saarbrücken</b>
<i>Recherchebasis: Titel aus Hildesheim</i>				
E-Commerce (65 Titel ges.)	10	5	9	2
Informationsmanagement (143 Titel ges.)	10	3	4	1
Information Retrieval (189 Titel ges.)	10	7	4	4
Maschinelle Übersetzung / Computerlinguistik (125 Titel ges.)	10	7	3	2
	<b>Hildesheim</b>	<b>Konstanz</b>	<b>Regensburg</b>	<b>Saarbrücken</b>

### 3 Evaluierung ausgewählter Bibliotheken mit dem Sachgebiet Informationswissenschaft

Mensch-Maschine-Interaktion (62 Titel ges.)	10	6	4	3
<i>Recherchebasis: Titel aus Konstanz</i>				
Informationspsychologie (103 Titel ges.)	3	10	1	0
Informationsrecht (25 Titel ges.)	1	11	4	1
Informationstechnologie (Software) (385 Titel ges.)	1	12	3	0
Informationsverhalten (7 Titel ges.)	1	7	6	0
Klassifizieren (16 Titel ges.)	1	10	1	0
<i>Recherchebasis: Titel aus Regensburg</i>				
Informationsnutzung (31 Titel ges.)	5	8	10	2
Informationsspeicherung (7 Titel ges.)	1	6	7	0
Informationssysteme (18 Titel ges.)	2	4	10	0
Informationsvermittlung (39 Titel ges.)	3	10	14	2
Klassifikationssysteme (16 Titel ges.)	1	6	9	1
<i>Recherchebasis: Titel aus Saarbrücken</i>				
Datenbanksoftware (7 Titel ges.)	1	2	2	7
Elektronisches Publizieren (42 Titel ges.)	4	8	9	13
Evaluation von Informationssystemen (15 Titel ges.)	1	6	4	12
Expertensysteme (27 Titel ges.)	4	4	5	9
Informationsnetze (10 Titel ges.)	1	2	1	9

Tabelle 3: Treffermengen der informationswissenschaftlichen Literaturrecherche  
in vier Bibliotheken

Bei einer Auswertung der Tabellenwerte muss selbstverständlich bedacht werden, dass die ermittelten Zahlen nur Tendenzen wiedergeben und keinen lückenlosen Vergleich der Teilbestände darstellen können. An der Gesamtmenge der Titel, die für die Recherche aus den einzelnen Teilgebieten der Klassifikationen ausgewählt wurden, lässt sich allerdings ablesen, dass man in 13 von 20

Fällen von einer repräsentativen Anzahl sprechen kann, die auch die entsprechenden Treffer bei der Suche in anderen Bibliotheken erwarten lässt.

Aus 5 der 20 Unterrubriken wurden alle bzw. fast alle Titel für die Literatursuche entnommen, aus vier von ihnen etwa die Hälfte der Bücher und aus vier ungefähr ein Drittel. Bei den fünf Themengebieten mit Hildesheimer Titeln sieht die Situation etwas anders aus: Da es sich dort um Lehrgebiete der Hildesheimer Informationswissenschaft handelt, findet sich dort eine sehr gute Ausstattung mit Literatur (65 – 189 Titel). Allerdings wurden hier für die Recherche fast ausschließlich aktuelle Werke der letzten Jahre ausgewählt, so dass auch dort eine große Wahrscheinlichkeit besteht, sie an anderen Standorten wiederzufinden. Bei den Konstanzer Gliederungspunkten „Informationspsychologie“ (10 von 103 Titeln) und „Informationstechnologie (Software)“ (12 von 385 Titel ges.) stellt die gesuchte Anzahl der Titel nur einen kleineren Ausschnitt der Gesamtmenge dieses Zweiges der Informationswissenschaft dar. Die Ursache dafür ist sicherlich auch in der deutlich besseren Ausstattung der Konstanzer Universitätsbibliothek zu suchen. Genau wie bei der Auswahl der Literatur in der Regensburger UB wurden auch in Konstanz nicht nur Titel mit einer informationswissenschaftlichen Signatur (AN bzw. kid) berücksichtigt, sondern ebenso solche, die mit einer entsprechenden zusätzlichen Notation bzw. Systemstelle versehen worden sind. Dadurch erweiterte sich der Kreis der in Frage kommenden Bücher, was aus Gründen der Vollständigkeit durchaus wünschenswert ist und die Aussagekraft des Ergebnisses erhöht. Insgesamt kann man also davon sprechen, dass die Ergebniszahlen Rückschlüsse auf die Literaturversorgung in den einzelnen Teilgebieten zulassen und Vergleiche zwischen ihnen somit möglich sind.

Sofort ins Auge fällt die insgesamt schlechte informationswissenschaftliche Ausstattung der UB Saarbrücken. Bezogen auf Anfragen mit Literatur anderer Standorte der Informationswissenschaft kann man lediglich im Bereich „Information Retrieval“ von einer durchschnittlich guten Bereitstellung von Literatur sprechen (vier von zehn). Vergleiche innerhalb von Themenbereichen, bei de-

nen die eigene Literatur als Basis zugrunde gelegt wurde, sind selbstverständlich nicht möglich.

Für die Universitätsbibliothek Regensburg stellt sich die Situation deutlich besser dar. In 4 von 15 Themengebieten – die in Regensburg ausgewählten Themengebiete würden wieder die Aussage verfälschen – kann man eine mangelhaften Ausstattung vermuten (ein oder zwei Titel), acht Teilbereiche bieten eine durchschnittliche Versorgung mit informationswissenschaftlicher Literatur an (drei bis fünf Titel) und in drei Bereichen ist ein gutes Angebot vorzufinden (sechs bis neun Titel).

Etwas schlechter, aber immer noch besser als für die Saarbrücker Institutsbibliothek, fällt die Einschätzung für den Hildesheimer Standort aus. Die vier Bereiche „Informationspsychologie“, „Informationsnutzung“, „Elektronisches Publizieren“ und „Expertensysteme“ weisen immerhin eine durchschnittliche Ausstattung auf (drei bis fünf Titel). Die restlichen elf Gebiete sind eher schlecht mit der notwendigen Literatur bestückt.

Wie schon das Ergebnis der Recherche für die Gesamtausstattung vermuten ließ, ist auch die Literaturversorgung in den einzelnen Themengebieten der Informationswissenschaft in der Konstanzer Universitätsbibliothek am komfortabelsten. Bei 9 von 15 Themenbereichen anderer Standorte weist Konstanz einen guten bis sehr guten Bestand auf (6 bis 10 gefundene Titel von 7 bis 14 recherchierten Titeln), bei vier einen durchschnittlichen (drei bis fünf Titel von neun bis zehn Titeln). Nur in zwei Bereichen („Datenbanksoftware“ und „Informationsnetze“) ist das Angebot schlecht (zwei Titel von sieben bzw. von neun Titeln). Bei der Auswertung der Literaturrecherche nach Konstanzer Titeln fällt auf, dass sich in den anderen Bibliotheken nur wenige dieser Bücher finden ließen, nur 20 von 50 Titeln wurden überhaupt gefunden. Dies kann zum einen daran liegen, dass sich in den umfangreichen Konstanzer Teilbereichen „Informationspsychologie“ und „Informationstechnologie (Software)“ viele Titel finden, die die anderen Bibliotheken nicht in ihrem Bestand führen und so die Wahrscheinlichkeit eines Suchtreffers sinkt. Dafür spricht die Tatsache, dass das Ergebnis bei den weniger umfangreichen Themengebieten „Informationsrecht“,

„Informationsverhalten“ und „Klassifizieren“, zumindest bezogen auf die Regensburger Bibliothek, besser ausfällt. In jedem Fall liegt der Schluss nahe, dass Konstanz in diesen Sachgebieten eine größere Auswahl an Literatur als die anderen Standorte vorzuweisen hat.

Unabhängig von der Menge der gefundenen Literatur ist noch ein Vergleich der Aktualität der klassifizierten Literatur von Interesse. Bei der Recherche war auffällig, dass im Gegensatz zu den anderen Standorten in Regensburg viele Werke älteren Erscheinungsdatums zu finden sind. Bei der zufälligen Auswahl aus den Systemstellen der Regensburger Klassifikation befanden sich unter 50 Büchern 5 Titel, die in den 60ziger Jahren publiziert worden waren und 20 Titel, deren Veröffentlichungsdatum zwischen 1970 und 1980 lag. Ein Teil dieser Bücher fand sich auch an den anderen Standorten wieder. Die ausgewählte Literatur war dort aber durchweg jüngeren Datums, was die Vermutung nahe legt, dass sich in Regensburg ein größerer Anteil an älterer informationswissenschaftlicher Literatur befindet.

### **3.1.3 Erfassung der Titel in den einzelnen Klassifikationen**

Von Interesse bei der Auswertung ist auch die Frage, wieviele der gefundenen Titel in den einzelnen informationswissenschaftlichen Klassifikationen erfasst sind, bzw. an welcher Stelle der entsprechenden Signatursystematik sie sonst geführt werden. Für die Bücher, die sich in der Klassifikation wiederfinden, ist es weiter wichtig, nach der Einordnung in die vorliegenden Themengruppen zu forschen.

#### **3.1.3.1 Saarbrücken**

Die im Saarbrücker Virtuellen Bibliotheksinformationssystem (ViBI) gefundenen 18 Titel der anderen Standorte sind ausnahmslos mit einer informationswissenschaftlichen Signatur versehen, da das ViBI auf die Institutsbibliothek der Informationswissenschaft zugreift. Bezogen auf die einzelnen Themengebiete der Suche finden sich die Saarbrücker Bücher an folgenden Stellen der institutseigenen Klassifikation:

3 Evaluierung ausgewählter Bibliotheken mit dem Sachgebiet  
Informationswissenschaft

<b>Themengebiete</b>	<b>Saarbrücker Systemstellen</b>
E-Commerce (HI) – 2 Titel	<ul style="list-style-type: none"> <li>• Informationstechnik ; EDV – Allgemein (1)</li> <li>• Informationstechnik ; EDV – Elektronisches Publizieren (1)</li> </ul>
Informationsmanagement (HI) – 1 Titel	<ul style="list-style-type: none"> <li>• Information und Dokumentation – Fachkommunikations- und Informationsprozesse, Informationsmanagement – Verfahren der Informationsgewinnung (1)</li> </ul>
Information Retrieval (HI) – 4 Titel	<ul style="list-style-type: none"> <li>• Informationstechnik ; EDV – Dialogsysteme – Information-Retrieval-Systeme (2)</li> <li>• Information und Dokumentation – Fachkommunikations- und Informationsprozesse, Informationsmanagement – Verfahren der Informationsgewinnung (1)</li> <li>• Information und Dokumentation – Fachkommunikations- und Informationsprozesse, Informationsmanagement – Allgemein (1)</li> </ul>
Maschinelle Übersetzung/Computerlinguistik (HI) – 2 Titel	<ul style="list-style-type: none"> <li>• Informationslinguistik ; Sprache und Texte in Informationsprozessen – Computerlinguistik – Maschinelle Übersetzung (1)</li> <li>• Informationslinguistik ; Sprache und Texte in Informationsprozessen – Computerlinguistik – Grammatiken, Parsing (1)</li> </ul>
Mensch-Maschine-Interaktion (HI) – 3 Titel	<ul style="list-style-type: none"> <li>• Informationstechnik ; EDV – Software-Ergonomie – Mensch-Maschine-Schnittstelle (2)</li> <li>• Allgemeines – Messen / Tagungen (1)</li> </ul>
Informationsrecht (KN) – 1 Titel	<ul style="list-style-type: none"> <li>• Informationswissenschaft – Gesellschaftspolitische Aspekte / Informationspolitik – Informationsrecht (1)</li> </ul>
Informationsnutzung (R) – 2 Titel	<ul style="list-style-type: none"> <li>• Allgemeines – Messen / Tagungen (1)</li> <li>• Information und Dokumentation – Fachkommunikations- und Informationsprozesse, Informationsmanagement – Verfahren der Informationsgewinnung (1)</li> </ul>
Informationsvermittlung (R) – 2 Titel	<ul style="list-style-type: none"> <li>• Information und Dokumentation – Informations- und Dokumentationsstellen – Bibliotheken und Bibliotheksnetze (1)</li> <li>• Information und Dokumentation – Informationssysteme – Allgemein (1)</li> </ul>
Klassifikationssysteme (R) – 1 Titel	<ul style="list-style-type: none"> <li>• Information und Dokumentation – Repräsentation von codiertem Wissen – Ordnung, Klassifikation, Thesaurus (1)</li> </ul>

Tabelle 4: Einordnung recherchierter Saarbrücker Titel in der dortigen Klassifikation



(Die Kürzel in Klammern (HI, KN, R, SB) hinter den Themengebieten stehen für die einzelnen Bibliotheksorte; die Zahlen in Klammern in der rechten Spalte zeigen die Anzahl der unter der jeweiligen Systemstelle platzierten Titel an.)

Durch die relativ geringe Anzahl an gefundenen Büchern ist die Beurteilung einer themenbezogenen Erfassung kaum möglich. Sie kann nur eingeschränkt für die wenigen Themen erfolgen, wo mehrere Titel im Trefferergebnis erscheinen. Dort kann man feststellen, dass diese in relativer Nähe zueinander erfasst worden sind, d.h. sie stehen unter derselben Oberkategorie oder finden sich in derselben Unterrubrik wieder. Für das Themengebiet Information Retrieval trifft dies allerdings nicht zu, wobei hier erneut eine Problematik deutlich wird, die durch eine virtuelle Signatur beseitigt werden könnte: Die Aufnahme der Titel ist sowohl unter der Rubrik „Information-Retrieval-Systeme“ als auch unter der Rubrik „Verfahren der Informationsgewinnung“ gerechtfertigt.

### 3.1.3.2 Regensburg

Im Gegensatz zum ViBI ist nur ein geringer Teil der Regensburger informationswissenschaftlichen Werke an der entsprechenden Stelle der Regensburger Verbundklassifikation, nämlich unter AN 90000 – AN 99000 (Informationswissenschaft) erfasst. Bei der Suche im Regensburger OPAC ergab sich, dass von 110 Regensburger Titeln der Literaturrecherche nur 34 mit der Notation AN 9XXXX (X steht für eine weitere Zahl) versehen sind. Insgesamt bietet sich bezogen auf die Einordnung der Bücher in die verschiedenen Notationen folgendes Bild (ein Titel kann mehrere Notationen besitzen):

<b>Themengebiete</b>	<b>Regensburger Systemstellen</b>
E-Commerce (HI) – 9 Titel	<ul style="list-style-type: none"><li>• Wirtschaftswissenschaften</li><li>• Informatik</li></ul>
Informationsmanagement (HI) – 4 Titel	<ul style="list-style-type: none"><li>• Wirtschaftswissenschaften</li></ul>
Information Retrieval (HI) – 4 Titel	<ul style="list-style-type: none"><li>• Allgemeine und vergleichende Sprach- und Literaturwissenschaft</li><li>• Informatik</li><li>• Informationsvermittlung, Informationsplanung (AN) (1)</li></ul>
Maschinelle Übersetzung / Computerlinguistik (HI) –	<ul style="list-style-type: none"><li>• Allgemeine und vergleichende Sprach- und Literaturwissenschaft</li></ul>

3 Evaluierung ausgewählter Bibliotheken mit dem Sachgebiet  
Informationswissenschaft

3 Titel	<ul style="list-style-type: none"> <li>• Informatik</li> </ul>
Mensch-Maschine-Interaktion (HI) – 4 Titel	<ul style="list-style-type: none"> <li>• Psychologie</li> <li>• Wirtschaftswissenschaften</li> <li>• Informatik</li> <li>• Allgemeine und vergleichende Sprach- und Literaturwissenschaft</li> </ul>
Informationspsychologie (KN) – 1 Titel	<ul style="list-style-type: none"> <li>• Psychologie</li> </ul>
Informationsrecht (KN) – 4 Titel	<ul style="list-style-type: none"> <li>• Rechtswissenschaft</li> <li>• Informatik</li> <li>• Rechtsfragen, Datenschutz (AN) (1)</li> </ul>
Informationstechnologie (Software) (KN) – 3 Titel	<ul style="list-style-type: none"> <li>• Wirtschaftswissenschaften</li> <li>• Informatik</li> </ul>
Informationsverhalten (KN) – 6 Titel	<ul style="list-style-type: none"> <li>• Psychologie</li> <li>• Soziologie</li> <li>• Wirtschaftswissenschaften</li> <li>• Informatik</li> <li>• Bibliographien, Sammelchriften – Tagungs- und Kongressberichte (AN) (2)</li> <li>• Medien- und Kommunikationswissenschaften, Kommunikationsdesign</li> </ul>
Klassifizieren (KN) – 1 Titel	<ul style="list-style-type: none"> <li>• Wirtschaftswissenschaften</li> </ul>
Datenbanksoftware (SB) – 2 Titel	<ul style="list-style-type: none"> <li>• Informatik</li> </ul>
Elektronisches Publizieren (SB) – 9 Titel	<ul style="list-style-type: none"> <li>• Informatik</li> <li>• Wirtschaftswissenschaften</li> <li>• Bibliothekswesen</li> <li>• Wissenschaftskunde und Wissenschaftsorganisation – Technik der wissenschaftlichen Arbeit</li> <li>• Medien- und Kommunikationswissenschaften, Kommunikationsdesign</li> </ul>
Evaluation Informationssysteme (SB) – 4 Titel	<ul style="list-style-type: none"> <li>• Wirtschaftswissenschaften</li> <li>• Psychologie</li> <li>• Informatik</li> <li>• Allgemeine und vergleichende Sprach- und Literaturwissenschaft</li> </ul>
Expertensysteme (SB) – 5 Titel	<ul style="list-style-type: none"> <li>• Informatik</li> </ul>
Informationsnetze (SB) – 1 Titel	<ul style="list-style-type: none"> <li>• Informatik</li> </ul>

Tabelle 5: Einordnung recherchierter Regensburger Titel in der dortigen Klassifikation

(Die Kürzel in Klammern (HI, KN, R, SB) hinter den Themengebieten stehen für die einzelnen Bibliotheksorte; die Zahlen in Klammern in der rechten Spalte zeigen die Anzahl der unter der jeweiligen informationswissenschaftlichen Sys-

temstelle platzierten Titel an. Diese sind auch durch den Zusatz „(AN)“ gekennzeichnet.)

Bei der Analyse der Suchtreffer fremder Standorte im Regensburger Web-OPAC zeigt sich, dass nur vier der gefundenen Titel mit einer Notation AN 9XXXX versehen sind, welche dieses Buch damit dem Bereich Informationswissenschaft zuordnet. Die restlichen Treffer tragen Notationen anderer Sachgebiete der Universitätsbibliothek Regensburg. Es findet sich also, abgesehen vom Thema „Informationsrecht“, keine Themen bezogene Entsprechung mit den anderen Bibliotheken bei der Zu- und Einordnung der informationswissenschaftlichen Literatur.

#### 3.1.3.3 Konstanz

Die Recherche im Konstanzer Ausleih- und Anfrage-System (Koala) ergab, dass von den 200 gesuchten Büchern sich 134 in der Konstanzer UB wiederfanden. Von diesen sind 98 mit einer eigenen Signatur in der Konstanzer Klassifikation kid erfasst. Die restlichen Titel tragen Signaturen anderer Sachgebiete der Bibliothek. Allerdings sind sie zumindest teilweise durch eine oder mehrere sogenannte Systemstellen (entsprechen den Hildesheimer Nebenstellen) versehen, so dass sie auf diesem Weg „virtuell“ im kid vertreten sind. Unglücklicherweise lässt sich anhand der Trefferanzeige im Koala nicht feststellen, ob eine weitere Systemstelle eingetragen ist, da dort lediglich die Signatur ausgewiesen wird (siehe Abbildung 5). Nur durch eine Abfrage über die Systemstelle ließe sich ermitteln, ob ein entsprechender Eintrag besteht.

Für einen möglichen Abgleich der Konstanzer mit der Hildesheimer informationswissenschaftlichen Literatur wurde nach der durchgeführten Literaturrecherche deutlich, dass immerhin 41 von 80 der in Hildesheim gefundenen Titel (51 Prozent) mit einer eigenen Signatur im kid aufgenommen sind (siehe beiliegende CD-ROM). Weitere 15 in Hildesheim geführte Titel finden sich unter Signaturen anderer Sachgebiete in der Konstanzer UB wieder. Die Suchanfragen nach einer bestimmten Notation an das Konstanzer Koala hatten ergeben, dass zumindestens ein Großteil dieser Bücher durch eine Nebenstelle im kid erfasst ist.

Addiert man also die beiden Werte auf, so kann man von bis zu 70 Prozent der Hildesheimer Titel (56 von 80) ausgehen, die im kid repräsentiert sind.

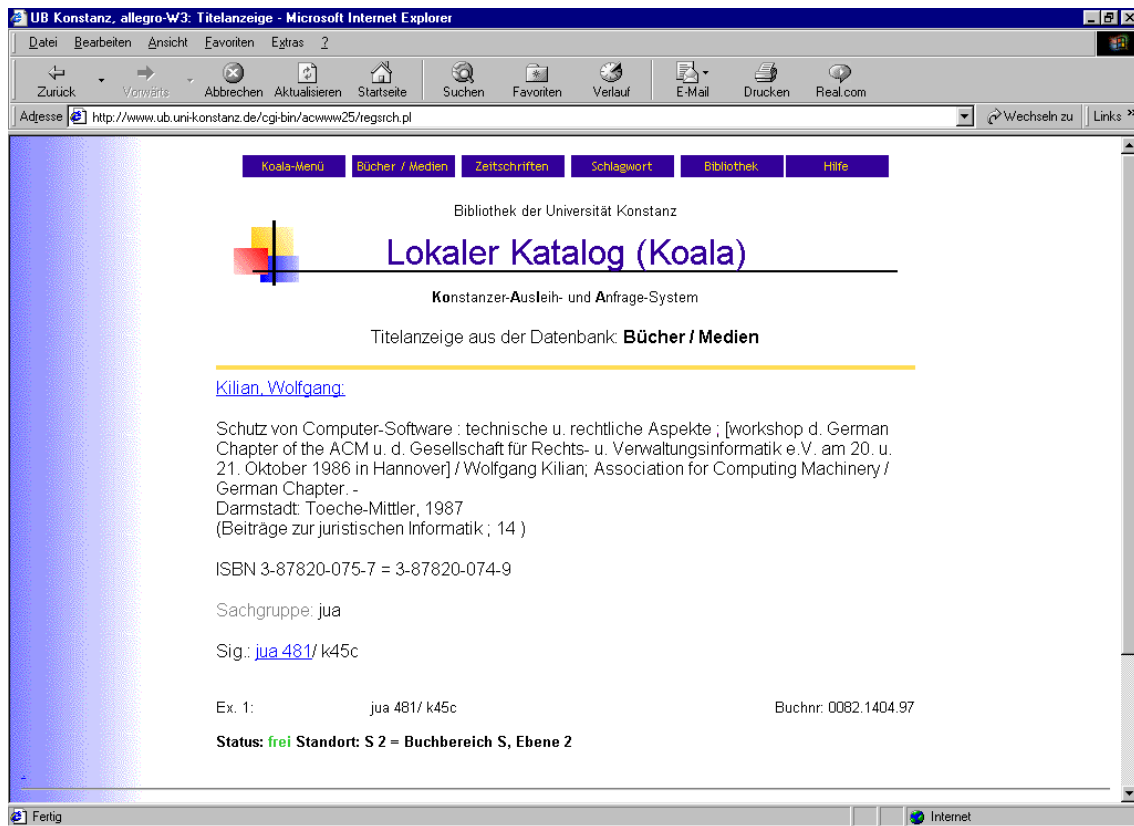


Abbildung 5: Beispiel für eine Trefferausgabe im Konstanzer Koala

Die Verteilung der Konstanzer Bücher über die einzelnen Sachgebiete der Bibliothek bzw. die einzelnen Themenbereiche der informationswissenschaftlichen Klassifikation hat folgendes Aussehen (Hier sind aufgrund der Vergabe von mehreren Notationen für denselben Titel Mehrfachnennungen möglich.):

Themengebiete	Konstanzer Systemstellen
E-Commerce (HI) – 5 Titel	<ul style="list-style-type: none"> <li>Wirtschaft, Management, Organisation – elektronischer Zahlungsverkehr (kid) (1)</li> <li>Internet, WWW = World Wide Web (kid) (1)</li> <li>Wirtschaftswissenschaften</li> </ul>
Informationsmanagement (HI) – 3 Titel	<ul style="list-style-type: none"> <li>Wirtschaft, Management, Organisation (kid) (1)</li> <li>Bewertungsverfahren, Informationsqualität (kid) (1)</li> <li>Wirtschaftswissenschaften</li> </ul>
Information Retrieval (HI) – 7 Titel	<ul style="list-style-type: none"> <li>Statistik</li> <li>Informationstechnologie – Software - Zugriffsverfahren, information retrieval, Suchverfahren, downloading (kid) (4)</li> </ul>

3 Evaluierung ausgewählter Bibliotheken mit dem Sachgebiet  
Informationswissenschaft

	<ul style="list-style-type: none"> <li>• Informationstechnologie – Software – Datenmodelle (kid) (1)</li> <li>• Bewertungsverfahren, Informationsqualität – Messverfahren (kid) (1)</li> </ul>
Maschinelle Übersetzung / Computerlinguistik (HI) – 7 Titel	<ul style="list-style-type: none"> <li>• Sprachwissenschaft</li> <li>• Sprachwissenschaft, Linguistik (kid) (1)</li> <li>• Informationslinguistik, Wissensrepräsentation – Sprachanalyse- und Sprachsynthese-Systeme, parsing (kid) (1)</li> <li>• Theoretische Informatik – Allgemeines (kid) (1)</li> </ul>
Mensch-Maschine-Interaktion (HI) – 6 Titel	<ul style="list-style-type: none"> <li>• Rechenzentren: Betriebssysteme, Verwaltung – Kommandosprachen, Benutzeroberflächen, user interface, Online-Hilfe, interaktive Systeme (kid) (2)</li> <li>• Kybernetik – Automat und Mensch, usability, Software-Ergonomie (kid) (2)</li> <li>• Kybernetik (kid)(1)</li> <li>• Frage-Antwort-Systeme, Expertensysteme (kid) (1)</li> </ul>
Informationsnutzung (R) – 8 Titel	<ul style="list-style-type: none"> <li>• Internet, WWW = World Wide Web (kid) (1)</li> <li>• Dokumentationssysteme (kid) (1)</li> <li>• Informationstechnologie – Software - Zugriffsverfahren, information retrieval, Suchverfahren, downloading (kid) (2)</li> <li>• Dokumentationstypen (kid) (1)</li> <li>• Informationsliteratur</li> <li>• Buch- und Bibliothekswesen</li> <li>• Datenbankführer: ODIN, Online-Führer (kid) (1)</li> </ul>
Informationsspeicherung (R) – 6 Titel	<ul style="list-style-type: none"> <li>• Informationsspeicherung (kid) (1)</li> <li>• Informations-Management, Wirtschaftsinformatik – Rationalisierung und Automatisierung in Büro und Verwaltung, office automation, Bürokommunikation, Workflow-Management (kid) (1)</li> <li>• Datenspeicher (kid) (1)</li> <li>• Buch- und Bibliothekswesen</li> <li>• Lehrbuchsammlung</li> </ul>
Informationssysteme (R) – 4 Titel	<ul style="list-style-type: none"> <li>• Buch- und Bibliothekswesen</li> <li>• Dokumentations- und Informations-Prozesse (kid) (1)</li> <li>• Informations-Management, Wirtschaftsinformatik – Planung, Organisation, Arbeitsablauf, Ergonomie (kid) (1)</li> </ul>
Informationsvermittlung (R) – 10 Titel	<ul style="list-style-type: none"> <li>• Bibliotheken, Buchhandel, digital libraries, Inetbib (kid) (1)</li> <li>• Informationslinguistik, Wissensrepräsentation (kid) (1)</li> <li>• Informations-Management, Wirtschaftsinforma-</li> </ul>

### 3 Evaluierung ausgewählter Bibliotheken mit dem Sachgebiet Informationswissenschaft

	<ul style="list-style-type: none"> <li>• tik – Rationalisierung und Automatisierung in Büro und Verwaltung, office automation, Büro-kommunikation, Workflow-Management (kid) (1)</li> <li>• Informations-Management, Wirtschaftsinformatik - Marketing, information broker (kid) (3)</li> <li>• Dokumentation, Informationstechnologie (kid) (1)</li> <li>• Theoretische Informatik – Allgemeines (kid) (1)</li> <li>• Buch- und Bibliothekswesen</li> </ul>
Klassifikationssysteme (R) – 6 Titel	<ul style="list-style-type: none"> <li>• Buch- und Bibliothekswesen</li> </ul>
Datenbanksoftware (SB) – 2 Titel	<ul style="list-style-type: none"> <li>• Datenbanken – spezielle Datenbanksysteme und Abfragesprachen (kid) (2)</li> </ul>
Elektronisches Publizieren (SB) – 8 Titel	<ul style="list-style-type: none"> <li>• HTML, CGI, XML WML (kid) (1)</li> <li>• Web Publishing, Edieren von Internet-Seiten Software (kid) (1)</li> <li>• Wirtschaft, Management, Organisation – elektronischer Zahlungsverkehr (e-commerce ; e-business) (kid) (1)</li> <li>• Informations-Dienstleistungen – Verlagswesen: publishing on demand, elektronisches Publizieren, FORMEX (kid) (1)</li> <li>• Verwaltungslehre</li> <li>• Pädagogik</li> </ul>
Evaluation Informationssysteme (SB) – 5 Titel	<ul style="list-style-type: none"> <li>• Kybernetik – Automat und Mensch, usability, Software-Ergonomie (kid) (3)</li> <li>• Computergraphik, Grafik-Formate, GKS = graphisches Kernsystem – Cliparts, Icons (kid) (1)</li> <li>• Informationslinguistik, Wissensrepräsentation – Sprachanalyse- und Sprachsynthese-Systeme, parsing (kid) (1)</li> </ul>
Expertensysteme (SB) – 4 Titel	<ul style="list-style-type: none"> <li>• soziologischer Aspekt der Information, Gefahren, Risiken (kid) (1)</li> <li>• Informationslinguistik, Wissensrepräsentation (kid) (1)</li> <li>• Frage-Antwort-Systeme, Expertensysteme (kid) (1)</li> <li>• Software für Expertensysteme: Babylon, Expertis, DEDUC (kid) (1)</li> </ul>
Informationsnetze (SB) – 2 Titel	<ul style="list-style-type: none"> <li>• Kommunikationstechnologie (kid) (1)</li> <li>• Kommunikationstechnologie – Breitbandnetze, ISDN (kid) (1)</li> </ul>

Tabelle 6: Einordnung recherchierter Regensburger Titel in der dortigen  
Klassifikation

(Die Kürzel in Klammern (HI, KN, R, SB) hinter den Themengebieten stehen für die einzelnen Bibliotheksorte; die Zahlen in Klammern in der rechten Spalte zeigen die Anzahl der unter der jeweiligen Systemstelle platzierten Titel an. Diese sind durch den Zusatz „(kid)“ gekennzeichnet.)

Im Vergleich zur Regensburger Universitätsbibliothek fällt auf, dass in Konstanz die informationswissenschaftlichen Bücher über eine geringe Anzahl anderer Sachgebiete verstreut stehen und in größerem Maße in der dafür bestimmten Klassifikation erfasst sind. Allerdings stellt man fest, dass für manche Themengebiete eines anderen Standortes nicht immer genaue oder sehr ähnlich geartete Entsprechungen in der Konstanzer Klassifikation existieren, sondern dass sich die Literatur dieses Zweiges der Informationswissenschaft nach anderen Klassifizierungsgesichtspunkten in Konstanz über mehrere Unterrubriken verteilt (z.B. „Mensch-Maschine-Interaktion“ oder „Informationsnutzung“). Auch wenn ähnliche Themenbereiche bestehen, verteilen sich die entsprechenden Titel über mehrere Rubriken der Konstanzer Klassifikation (z.B. „Expertensysteme“, „Elektronisches Publizieren“). Lediglich in den Bereichen „Information Retrieval“ und „Informationsvermittlung“ findet man eine Trefferkonzentration in ähnlich bezeichneten Unterpunkten der Klassifikation.

Bei diesem Resultat muss sicherlich auch berücksichtigt werden, dass viele der gewählten Zweige der Informationswissenschaft eine breitere thematische Auffächerung zulassen. Auch weist die Konstanzer Klassifikation im Vergleich zu den beiden anderen untersuchten Standorten eine feinere und detailliertere Gliederung auf, die eine solche Streuung der Literatur begünstigt. Trotzdem liefert diese stichpunktartige Untersuchung ein deutliches Indiz für die eingeschränkte Vergleichbarkeit zwischen verschiedenen Klassifikationsrubriken, die sich aus den unterschiedlichen Aspekten und Kriterien bei der Klassifizierung ergibt.

#### *3.1.3.4 Thematische Kongruenzen zwischen den Klassifikationen*

Ein weiterer Aspekt der Untersuchung war, herauszufinden, in welchen Teilbereichen Überschneidungen zwischen den Klassifikationen bestehen. Für den

Benutzer kann dies von Bedeutung sein, wenn er nämlich mit dem Suchergebnis in der Bibliothek vor Ort nicht zufrieden ist und er weitere Standorte für seine Recherche mit heranziehen möchte. Er kann dann auf die seiner Suchanfrage entsprechenden Positionen in anderen Klassifikationen verwiesen werden, ohne dass er sich mit den bestehenden dortigen Strukturen des Literaturangebotes vertraut machen muss. Aus den Analysen der Kapitel 3.1.3.1 bis 3.1.3.3 resultiert allerdings, dass die Erstellung von Kongruenzen zwischen den Klassifikationen der drei untersuchten Standorte nicht unproblematisch ist.

Für die Universitätsbibliothek Saarbrücken muss man die geringe Anzahl der dort gefundenen Büchern bedenken. Zwar scheinen die Ergebnisse auf eine einheitliche und übersichtliche Erfassung in Bezug auf die Themengebiete hinzudeuten, für eine Beurteilung ist die Ausstattung der Institutsbibliothek – und damit das Suchergebnis – aber nicht ausreichend umfangreich.

Die Problematik in Regensburg besteht im geringen Grad der Erfassung von informationswissenschaftlicher Literatur in der dafür vorgesehenen Klassifikation. Die Bücher stehen verstreut über zahlreiche andere Sachgebiete der Regensburger Verbundklassifikation, so dass eventuell einander ähnliche Themengebiete gefunden werden können, die in Regensburg dort auffindbaren Titel aber nur einen Ausschnitt des Literaturangebotes darstellen, welches in der gesamten Bibliothek zu diesem Teilgebiet der Informationswissenschaft zu finden ist.

Für die Konstanzer Universitätsbibliothek gilt es, die detaillierte Gliederung der Klassifikation zu beachten. Eine wie auch immer geartete „Verlinkung“ zwischen verschiedenen Klassifikationsrubriken kann wohl nur auf unterster Ebene des KID erfolgen.

### **3.2 Evaluierung der informationswissenschaftlichen Klassifikationen**

Wie im Kapitel 2.2.1 beschrieben, waren aus den aufgeführten Gründen die Bibliotheken der Standorte Konstanz, Regensburg und Saarbrücken mit ihren



informationswissenschaftlichen Systematiken als Grundlage für die weiteren Arbeiten ausgewählt worden. Neben dieser ersten Entscheidung sollte in einem weiteren Schritt untersucht werden, wie die Beschaffenheit der einzelnen ausgesuchten Klassifikationen ist, welche Vor- und Nachteile sie aufweisen. Diese Begutachtung war insofern von Bedeutung, als sie das Vorgehen beim Aufbau einer eigenen Klassifikation für den informationswissenschaftlichen Bücherbestand der UB Hildesheim erleichtern sollte.

Um sich für die anstehende Arbeit die notwendige Übersicht zu verschaffen, wurde zuerst einmal von jeder der drei Klassifikationen ein Schaubild erstellt, in welchem sich die einzelnen Teilgebiete wiederfanden (siehe Anlagen 2, 3, 4). Dabei konnten für die Standorte Regensburg und Saarbrücken alle Ober- und Unterrubriken berücksichtigt werden. Aus Platzgründen fanden aus der Konstanzer Klassifikation nur die Überbegriffe der Gliederungen Eingang in das Schaubild, woraus man wiederum auf die schon erwähnte, sehr detaillierte Auffächerung dieser Systematik schließen kann.

Neben dem unterschiedlichen Umfang und Detaillierungsgrad ist auch das Entstehungsdatum bzw. das Datum der letzten Aktualisierung von Interesse. Laut einer E-Mail vom 22.04.2002 von Frau Barbara Leiwesmeyer (RVK-Verbundkoordinatorin der Universitätsbibliothek Regensburg) fand die letztmalige Überarbeitung der Regensburger Verbundklassifikation, von der die informationswissenschaftliche Systematik einen Auszug darstellt, im März 2000 statt. Das letzte Update des Klassifikationsschemas der Saarbrücker Institutsbibliothek für Informationswissenschaft fand im November 1999 statt (vgl. Universität Saarbrücken). Am aktuellsten präsentiert sich das Konstanzer kid, auf dessen Seiten man den November 2000 als Datum der letzten Aktualisierung findet (vgl. Universitätsbibliothek Konstanz).

Der unterschiedliche Aktualisierungsgrad korreliert allerdings nicht ganz mit den Inhalten der Klassifikationen. In der Regensburger Systematik (siehe Anlage 2) findet man viele Teilgebiete, die sich mit der Dokumentenvorbereitung und der Dokumentensuche, sowie den dafür notwendigen technischen Verfahren beschäftigen. Ein starker Schwerpunkt liegt dort also auf dem Information Retrie-

val, bezogen auf Textdokumente. Drei der acht Oberrubriken („Grundlagen / Theorien“, „Thesauri / Begriffsverzeichnisse“, „Informationspraxis“) befassen sich direkt mit der Dokumentenverarbeitung, zwei („Informationsvermittlung / Informationsplanung“ und „Fachliche Information und Dokumentation“) mit der Bereitstellung der Dokumente und damit verbundenen Teilaspekten. Die Betonung dieser älteren Teilgebiete der Informationswissenschaft könnte auf das frühe Entstehungsdatum (lt. Leiwesmeyer 1975) der A-Systematik im RVK (in welcher die Informationswissenschaft untergebracht ist) zurückzuführen sein und darauf, dass sie offenbar noch nicht ausreichend an die neuesten Entwicklungen in dieser wissenschaftlichen Disziplin angepasst wurde. Neuere Wissensgebiete der Informationswissenschaft, wie z.B. „Elektronisches Publizieren“, „Informationsvisualisierung“ oder „Software Ergonomie“, finden sich dort beispielsweise nicht.

Auffällig in der Regensburger Klassifikation sind die vorhandenen Redundanzen. In den Kategorien „Thesauri, Begriffsverzeichnisse“ und „Fachliche Information und Dokumentation“ findet man abgesehen von drei zusätzlichen Gliederungspunkten unter der letztgenannten Rubrik gleich lautende Themengebiete. Obwohl die inhaltliche Unterscheidung nachvollziehbar ist, ist es fraglich, ob man diese Einteilung als sinnvoll bezeichnen kann. Es ist meiner Meinung nach generell für einen Benutzer verwirrend, in einer Klassifikation auf der gleichen Ebene dieselben Bezeichnungen zu finden; überhaupt versuchen Bibliothekare bei der Erstellung von Klassifikationen, Redundanzen zu vermeiden (siehe Gesprächsprotokoll Müller / Pretz v. 17.04.2002, Anlage 21). Eine schnelle Orientierung wird so erschwert. Auch die Verwendung der Begriffe „Klassifikation“, „Klassifikationsvergleich“ und „Klassifikationssysteme“ auf der gleichen Ebene der Systematik dient nach meiner Ansicht nicht der Übersichtlichkeit.

Die Saarbrücker Klassifikation (siehe Anlage 3) zeigt sich in dieser Hinsicht zeitgemäßer und aktueller und hat einige dieser neueren Teilgebiete in ihre Systematik aufgenommen. Neben den schon erwähnten Themengebieten „Elektronisches Publizieren“ und „Software Ergonomie“ (unter der Kategorie „Informationstechnik; EDV“) findet man Bereiche wie „Computerlinguistik“ (unter

„Informationslinguistik“), „Anwendungssoftware“, „Programmierung, Programmiersprachen“ und „Hypertext / Hypermedia / Multimedia“ (alle unter „Informations-technik; EDV“). Ein Großteil der auf die Dokumentenverarbeitung und –aufbereitung bezogenen Rubriken der Regensburger Klassifikation wurden unter den Bereichen „Repräsentation von codiertem Wissen“ und „Fachkommunikations- und Informationsprozesse, Informationsmanagement“ in der Kategorie „Information und Dokumentation“ zusammengefasst.

Etwas verwirrend ist, dass die Saarbrücker Systematik eine eigene Kategorie „Informationswissenschaft“ beinhaltet, was übrigens auch beim Konstanzer kid der Fall ist. Dort befindet sich Literatur zu den Themen „Allgemeines“, „Berufsbilder und Ausbildung“, „Forschungseinrichtungen und Forschungsprojekte“, „Gesellschaftspolitische Aspekte und Informationspolitik“ und „Kognitionstheorien“. Beim letztgenannten Punkt dieser Rubrik fällt der fehlende Bezug zu den ersten drei Kategorien auf. Sicherlich hätte man für diese Rubrik auch eine passendere Oberkategorie finden können; für einen Benutzer ist die gewählte Bezeichnung „Informationswissenschaft“ jedenfalls auf den ersten Blick irreführend und nicht hilfreich.

Unter der Kategorie „Allgemeines“ trifft man wiederum auf Themenbereiche, die auch schon unter „Informationswissenschaft“ auftauchen. Bei „Gesetze und Gesetzesentwürfe“, „Forschung“ und „Bildungsstätten“ ist die enge inhaltliche Verwandtschaft zu den unter „Informationswissenschaft“ geführten Büchern leicht erkennbar. Eine Zusammenfassung bietet sich hier an. Die anderen Punkte unter „Allgemeines“ umfassen im wesentlichen für die Informationswissenschaft relevante Nachschlagewerke und Schriften.

Abgesehen von geringen Abweichungen sind die Inhalte der beiden letztgenannten Saarbrücker Kategorien, „Informationswissenschaft“ und „Allgemeines“, identisch mit den Regensburger Kategorien „Bibliographien, Sammelschriften“, „Organisation, Beruf, Forschung“ und „Nachschlagewerke, Allgemeine Darstellungen“.

Wichtige Neuerungen im Vergleich zur Regensburger Klassifikation enthält, neben den schon genannten Themenbereichen, die Saarbrücker Kategorie „Telekommunikation“. Dort werden Bereiche behandelt, die sich mit dem Informationsaustausch und der Kommunikation in Datennetzen beschäftigen. Auch die Datensicherheit wird hier thematisiert. Eine weitere Kategorie „Publikumsinformation“ beinhaltet Literatur zu den Themengebieten „Medien“ und „Public Relations“.

Die nach dem angegebenen Datum des letzten Updates neueste der drei untersuchten Klassifikationen, das Konstanzer kid (siehe Anlage 4), unterscheidet sich von den anderen beiden nicht nur durch seine größere Ausführlichkeit. Der Themenbereich der Elektronischen Datenverarbeitung nimmt hier einen sehr breiten Raum ein. Allein 8 der 13 Hauptgruppen dieser Gliederung befassen sich im engeren Sinne mit EDV: „Datenverarbeitung“, „Programmierung“, „Programmiersprachen“, „Dienstprogramme“, „Rechenzentren“, „Hardware“, „Betriebssysteme“, „Anwendungsbereiche der EDV“. Im weiteren Sinne kann man auch noch den Bereich „Künstliche Intelligenz“, der im kid unter „Anwendungsbereiche der EDV“ eingeordnet ist, zu dieser Gruppe zählen.

Eine Erklärung für diese starke EDV-Lastigkeit ist wohl in der Entstehungsgeschichte des kid zu suchen. Der Name steht ja für „Kybernetik, Informatik, Datenverarbeitung und Informationswissenschaft“. Da der letzte Begriff keinen Eingang in die Abkürzung gefunden hat, kann man davon ausgehen, dass er als letzter Themenbereich zu dieser systematischen Gliederung hinzugefügt wurde. Dafür spricht auch, dass es eine eigene Hauptgruppe „Informationswissenschaft“ gibt.

Ein Vorteil daran ist sicherlich, dass so die einzelnen Themengebiete der EDV, von denen viele eine nicht zu vernachlässigende Bedeutung für die Informationswissenschaft haben, ausgiebiger als in den anderen Klassifikationen behandelt werden. Als Nachteil muss wohl aber gesehen werden, dass es dadurch zu einer starken Überrepräsentation dieses Bereiches kommt. Es finden sich einige Themen der Informatik und EDV, die nur geringe oder gar keine in-

Informationswissenschaftliche Relevanz haben (z.B. „Digitalelektronik“ und „Quanten-Computer“).

Ein weiterer Nachteil ist, dass es so zu gewissen Redundanzen kommt. Beispielsweise werden unter der Rubrik „Informationstechnologie“ in der Hauptgruppe „Informationswissenschaft“ die Unterrubriken „Hardware“ und „Software“ geführt. Es soll wohl hier der Schwerpunkt auf die informationswissenschaftliche Literatur gelegt werden. Diese Themenbezeichnungen finden sich allerdings schon an anderer Stelle in der Systematik, was für den Benutzer verwirrend ist. Außerdem lässt sich wohl bei diesen Bereichen nicht immer klar zwischen Büchern für Informatiker und Informationswissenschaftler unterscheiden. Benutzern, die nur jeweils eine der Rubriken durchforsten, entgeht so möglicherweise interessante und relevante Literatur. Eine Konzentration und Zusammenfassung der betroffenen Rubriken wäre im Sinne einer besseren Übersichtlichkeit sinnvoll. Ähnliches gilt für die folgenden Themenbereiche unter „Informationswissenschaft“: „Theorie und Methodologie“, „Datenbasen, Datenbanken“, „Ausbildung, Fortbildung“ und „Institutionen“.

Als Verbesserung im Vergleich zur Regensburger und Saarbrücker Klassifikation erweist sich die Erweiterung der Rubriken um neue Wissensgebiete. Dazu zählen beispielsweise die Themenbereiche „Informationsmanagement, Wirtschaftsinformatik“ und „Informationssoziologie“. In letzterem werden Bücher zu Themen wie „Benutzerforschung“, „Informationsbedürfnisse“ und „Informationsverhalten“ bereitgehalten.

Nachschlagewerke und andere allgemeine Schriften werden im kld unter den Hauptgruppen „Allgemeines“, „Bibliographien“ und „Festschriften“ aufgenommen. Theoretische Grundlagen zur Dokumentenerfassung und Informationsaufbereitung finden sich unter „Informationslinguistik, Wissensrepräsentation“, Literatur zur Informationstheorie wird in der gleichnamigen Rubrik zur Verfügung gestellt, die ansonsten hauptsächlich Literatur mit informatischem Schwerpunkt anbietet („Kombinatorik / Graphentheorie“, „Algorithmen / Komplexität“).

Als kurzes Fazit der Untersuchungen lässt sich folgendes festhalten:

- Die Regensburger Verbundklassifikation ist im Bereich der Informationswissenschaft in einigen Teilen veraltet und weist für den Benutzer irritierende Redundanzen auf.
- Die Saarbrücker Klassifikation präsentiert aktuellere Inhalte, allerdings verwirrt die Kategorie "Informationswissenschaft" und deren Inhalte. Auch in dieser Systematik finden sich einige Redundanzen.
- Das Konstanzer kid ist im Vergleich dazu sehr detailliert. Der Bereich EDV ist aber überpräsentiert und auch hier sind wieder die bestehenden Redundanzen zu bemängeln. Aufgrund seiner Ausführlichkeit beinhaltet das kid so gut wie alle Themenbereiche der anderen Klassifikationen. Aus diesem Grund ist es sehr geeignet, bei der Erstellung einer eigenen informationswissenschaftlichen Klassifikation für die UB Hildesheim als Basis zu dienen.

## **4 Der informationswissenschaftliche Bücherbestand in der UB Hildesheim**

### **4.1 Alphabetisches Register von informationswissenschaftlichen Begriffen und zugehörigen Notationen**

Um den Schwierigkeiten beim Auffinden informationswissenschaftlicher Literatur in der Universitätsbibliothek Hildesheim durch eine neue, zusätzliche Ordnungssystematik zu begegnen, wurde in einem ersten Schritt eine alphabetische Liste von informationswissenschaftlichen Begriffen und Themengebieten erstellt. Diese Liste enthält, genau wie das in Kapitel 2.2.1 beschriebene alphabetische Register der ULB Düsseldorf, Schlüsselwörter, die für den Aufbau einer informationswissenschaftlichen Klassifikation von Bedeutung sind. Die Liste entstand in Anlehnung an Bezeichnungen aus den Klassifikationen der Standorte Konstanz, Regensburg und Saarbrücken. Außerdem wurden die den Notationen der UB Hildesheim zugeordneten inhaltlichen Beschreibungen und die Oberbegriffe der Basisklassifikation für die Erstellung des Registers herangezogen. Die Themengebiete der Basisklassifikation sind den Notationen der lokalen Hildesheimer Systematik als Konkordanz zugeordnet, also z.B. in der Form: *CSC 410 Daten-Strukturen; Daten-Modellierung =54.62=*. In diesem Fall ist die Basisklassifikation 54.62 und steht für „Informatik / Datenstrukturen“.

Zuerst wurden sämtliche Notationsbeschreibungen der UB Hildesheim in einem gemeinsamen Verzeichnis abgespeichert. Sie standen im Internet auf den Seiten der Universitätsbibliothek zum Download zur Verfügung (vgl. Universitätsbibliothek Hildesheim<sup>2</sup>). Anschließend wurde eine ausführbare Datei (grep.exe), die im Rahmen des GNU-Projektes als freie Software angeboten wird (vgl. GNU Project, weitere Informationen dazu siehe Kapitel 4.3.1.1) in diesen Ordner gestellt; sie ermöglicht es, über einen aus der Konsole abgesetzten Suchbefehl alle Dateien des Verzeichnisses nach einem Suchbegriff zu durchforsten. Die Ausgabe erfolgte zeilenweise in eine neue Datei, die mit dem Namen des Suchbegriffs bezeichnet wurde. Der Konsolen-Aufruf lautete folgendermaßen: *grep Suchbegriff \*.txt >Suchbegriff.txt*.

Die Ergebnisdateien wurden intellektuell auf die Suchterme hin untersucht, und die entsprechenden Notationen auf ihre Relevanz für die neue Klassifikation geprüft. Als Suchergebnis zum Begriff „Dokument“ ergab sich als ein Beispiel die unten abgebildete Ergebnisdatei (siehe Abbildung 6). Es galt zu entscheiden, ob der gefundene Begriff in dem jeweiligen Kontext der Systemstelle einen ausreichend engen Bezug zur Informationswissenschaft aufweist, und somit eine Zuordnung der betreffenden Systemstelle zu der Klassifikation gerechtfertigt erschien. Auf die Notation *SOZ 501 Frauenleben / Dokumente, Materialien, Biographien* trifft diese Voraussetzung beispielsweise nicht zu. Die Ergebniszeile *CSC 773 Index-Erzeugung / Dokumentverarbeitung* lieferte hingegen ein relevantes Ergebnis, so dass CSC 773 den informationswissenschaftlichen Begriffen „Indexierung“ und „Dokumentverarbeitung“ zugeordnet werden konnte.

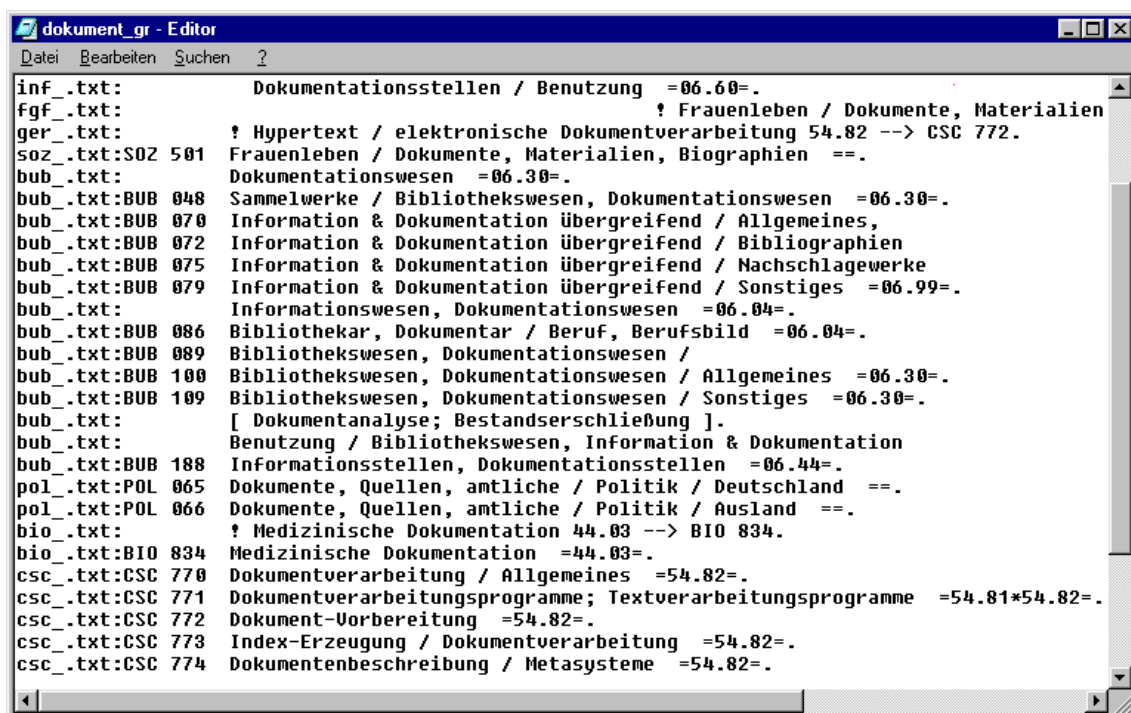


Abbildung 6: Ergebnis der Suchanfrage „Dokument“ an die Notationsbeschreibungen der UB Hildesheim

Alle zu einem Suchbegriff des alphabetischen Registers gefundenen Systemstellen wurden diesem in einer Tabelle zugeordnet (Auszug: siehe Anlage 5, komplette Tabelle: siehe beiliegende CD-ROM). Nicht bei allen Termen war die Recherche in den Notationen der UB Hildesheim erfolgreich. Teilweise exis-



tierten die Begriffe als Beschreibung nicht, teilweise war der informationswissenschaftliche Kontext nicht gegeben (z.B. Forschung).

### **4.2 Erfassung der relevanten Titelmenge in der UB Hildesheim über den Web-OPAC**

Damit für einen späteren Abgleich eine möglichst umfassende Menge aller informationswissenschaftlichen Bücher der Universitätsbibliothek Hildesheim vorliegt, wurde auf das erstellte alphabetische Register zurückgegriffen. Dieser Weg der Literaturbeschaffung wurde nach Rücksprache mit Herrn Edwin Pretz, dem EDV-Beauftragten und Fachreferenten für Informatik, Naturwissenschaften und Technik in der UB Hildesheim, beschritten.

Ursprünglich war über die Möglichkeiten einer Erfassung des Gesamtbestandes der UB Hildesheim nachgedacht worden. Dies hätte durch eine Datenbankabfrage an das Pica-System der UB Hildesheim realisiert werden müssen (Pica ist ein international operierendes Unternehmen im Bereich der Informationstechnologie mit Dienstleistungen für Bibliotheken und Kunden von Bibliotheken, vgl. Pica). Herr Pretz wandte gegen diese Idee ein, dass dafür ein umfangreiches SQL-Skript hätte geschrieben werden müssen, wofür eine genaue Kenntnis der Datenbankstruktur Voraussetzung gewesen wäre. Außerdem würde eine solche Abfrage eine starke Belastung des Datenbankservers darstellen, die im laufenden Betrieb nicht stattfinden könnte.

Anstatt dessen machte Herr Pretz den Vorschlag, sich die Daten durch einen gezielten Download relevanter Notationen über den Web-OPAC zu beschaffen. Hierzu erwies sich die OPC4 Testinstallation des GBV (Gemeinsamer Bibliotheksverbund) als geeignet. Sie stellt die Nachfolgeneration des alten Web-OPAC dar, welcher zum damaligen Zeitpunkt (Dezember 2001) noch in Benutzung war. Die neue Oberfläche befand sich damals noch in einer Testphase.

Der OPC4 weist im Vergleich zu seinem Vorgänger einige entscheidende Verbesserungen auf. Im Gegensatz zu vorher gibt es jetzt keine Trefferbegrenzung

mehr. Die Darstellung erfolgt zwar seitenweise mit jeweils nur zehn Titeln, über einen Download können aber bis zu 1000 Titel mit allen bibliographischen Angaben zur Ansicht gebracht werden. Als weiterer Vorteil ist die deutlich höhere Geschwindigkeit, mit der die Trefferliste ausgegeben wird. Der Grund dafür ist eine neue Suchmaschine mit der Bezeichnung PSI (Pica Search and Indexing).

Mehrere Aspekte mussten bei der Bestandserfassung über den OPC4 bedacht werden. Als Basis der Recherche wurden die Systemstellen der erstellten Liste informationswissenschaftlicher Begriffe zugrunde gelegt. Es mussten also die einzelnen Notationen als Suchbegriffe eingegeben werden. Ein Problem ergab sich aus der Tatsache, dass teilweise Titel durch ihre Signatur einer Notation zugeordnet sind, ohne dass die betreffende Notation noch einmal unter „Sachgruppen“ erscheint, wo sie als sogenannte „Nebenstelle“ dem Titel zugeordnet werden kann (siehe Abbildung 7). Bei einer reinen Suche nach Notationen der Sachgruppen würde man diese Titel also nicht in der Treffermenge wiederfinden. Notwendig war somit eine kombinierte Suchanfrage mit dem Booleschen Operator ODER, wodurch sowohl nach der Signatur als auch nach den Notationen der Nebenstellen gesucht werden konnte, ohne dass Dubletten anfielen. Eine beispielhafte Anfrage sah folgendermaßen aus: *kls csc 32? oder sgn csc 32?* (*kls* steht für Klassifikation und sucht in den Notationen der Sachgruppen, *sgn* für Signatur, das Fragezeichen steht für ein beliebiges Zeichen). Als Ergebnismenge erhielt man alle Titel, die eine Eintragung der Notationen *csc 320* bis *329* enthalten, entweder unter „Sachgruppen“ oder „Signatur“ oder unter beiden Rubriken (siehe Abbildung 7).

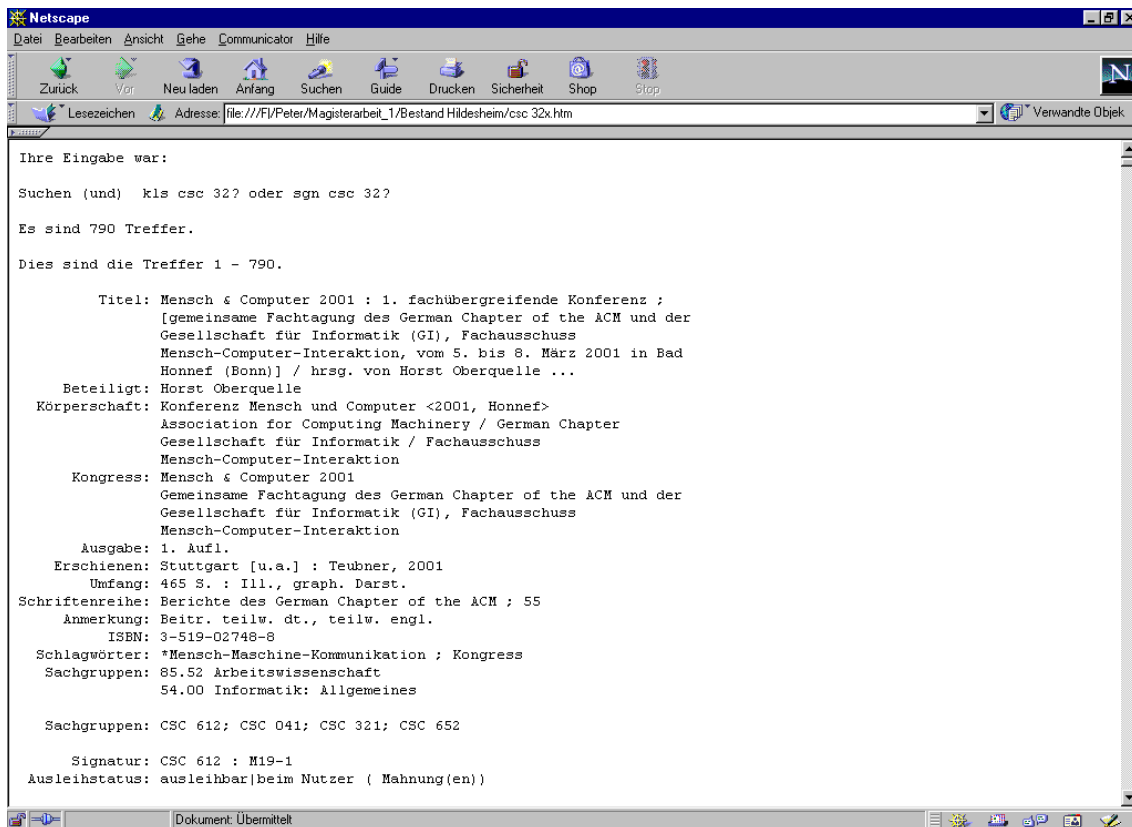


Abbildung 7: Beispiel des Ergebnisses einer Suchanfrage an den OPC4  
(Quelle: Universitätsbibliothek Hildesheim<sup>3</sup>)

Als nicht zu lösende Problematik muss man akzeptieren, dass in der Universitätsbibliothek Hildesheim noch nicht alle Titel mit ihren bibliographischen Angaben digital erfasst sind. Literatur aus dem Freihandbereich und dem Magazin, die vor 1990 angeschafft wurde, ist nur teilweise elektronisch in den Bestand eingepflegt. Erst wenn diese Bücher ausgeliehen werden, erfolgt eine digitale Aufnahme. Da es sich bei den informationswissenschaftlichen Büchern aber in der Regel um solche jüngeren Erscheinungsdatums handelt, fällt dieser Nachteil nicht allzu sehr ins Gewicht.

Eine weitere Problematik stellen Minimalsätze (auch als o.T. = ohne Titel gekennzeichnet) dar, die nur aus der Signatur und dem Ausleihstatus bestehen. Hier erfolgte aus Zeitgründen bei der elektronischen Erfassung der Dokumente der UB Hildesheim eine verkürzte Aufnahme in den digitalen Katalog. Da keine Titelangabe vorliegt, sind sie beim Abgleich nicht aufzufinden. Insgesamt han-

delt es sich aber um eine geringe Anzahl von Büchern, so dass dadurch die Bestandserfassung nur in zu vernachlässigender Weise beeinträchtigt wird.

Im Anschluss an das Herunterladen der einzelnen Ergebnismengen zu den oben beschriebenen Abfragen, wurden diese HTML-Dokumente in einer einzigen Datei zusammengefasst. Diese Textdatei mit einer Größe von ca. 13 MB sollte als Basis für den Abgleich mit anderen informationswissenschaftlichen Beständen dienen.

### **4.3 Abgleich mit dem informationswissenschaftlichen Bestand der UB Konstanz**

Nachdem die relevanten Titel der UB Hildesheim in einer Datei erfasst worden waren, fehlte für einen automatischen Abgleich das entsprechende Gegenstück, eine Datei mit einer möglichst großen Menge an informationswissenschaftlichen Titeln. Die Begutachtung und Evaluierung ausgewählter Bibliotheken (siehe Kapitel 3) hatte deutlich gemacht, dass im Vergleich zu den anderen untersuchten Büchereien das Konstanzer kid hier die weitaus umfangreichste Sammlung anbietet. Auch hatte die Literaturrecherche (siehe Kapitel 3.1) ergeben, dass sich danach ca. 70 Prozent der gefundenen Hildesheimer Titel (56 von insgesamt 80 Büchern) im kid wiederfinden lassen (Die Prozentzahl ist insofern unsicher als nicht klar ist, bei wie vielen der gefundenen Konstanzer Titel ohne kid-Signatur eine kid-Systemstelle eingetragen ist, die sie somit gewissermaßen zum erweiterten Kreis der Titel des kid gehören lässt. Eine eingetragene Systemstelle wird beim Suchergebnis eines Buches nicht angezeigt, lässt sich also so nicht nachweisen.)

Im Vergleich dazu sind die informationswissenschaftlichen Bestände der anderen Standorte für eine mengenmäßige Erfassung des Hildesheimer Bestandes als nicht geeignet zu bezeichnen. In der Regensburger UB ist nur ein kleiner Ausschnitt der informationswissenschaftlichen Titel mit einer entsprechenden Notation erfasst (insgesamt 970 Titel). Die meisten tragen die Signatur eines anderen Sachgebietes. Die durchgeführte Recherche hat gezeigt, dass von 80

Hildesheimer Titeln nur fünf mit einer informationswissenschaftlichen Notation der Regensburger UB versehen waren. Alle diese Titel findet man ebenfalls in Konstanz.

Der Bestand der Saarbrücker Institutsbibliothek ist mit 2.524 Titeln zwar etwas umfangreicher, aber immer noch zu klein, um bei einem Abgleich auf ein gutes Ergebnis hoffen zu dürfen. Hinzu kommt, dass von den 80 Hildesheimer Titeln zwar 26 in Saarbrücken geführt werden, diese bis auf fünf Titel aber auch alle in Konstanz zu finden waren.

Aus diesen Gründen wurde entschieden, das kid für den durchzuführenden automatischen Abgleich heranzuziehen. Der für das kid zuständige Konstanzer Fachreferent, Herr Gerhard Schmitz-Veltin, übersandte für diesen Zweck zwei Dateien. Die eine Datei enthielt alle Titel, die eine eigene kid-Signatur tragen, die zweite umfasste zusätzlich auch die Titel, bei denen eine kid-Notation als erste oder zweite Systemstelle eingetragen ist. Diese Titel besitzen eine Signatur aus einem anderen Sachbereich, beispielsweise aus Sprachwissenschaft, Recht oder Wirtschaft. Da die Anzahl solcher Bücher nach den Resultaten der Literaturrecherche immerhin eine Größenordnung von ca. 20 Prozent der in Hildesheim gefundenen Titel erreicht, sollten auch sie unbedingt mit in den Abgleich einbezogen werden. Stichpunktartige Untersuchungen ergaben, dass nicht alle Titel der ersten genannten Datei in der zweiten zu finden waren (was eigentlich der Fall hätte sein müssen), weshalb für den Abgleich beide genannten Dateien verwendet wurden.

Die zwei Dateien hatten ein unterschiedliches Aussehen. Während in der ersten Datei die Titel des kid in abgebrochener Form mit Signatur, sowie zugehörigem Autor und Erscheinungsjahr vorlagen, fanden sich in der zweiten Datei neben dem kompletten Titel der Autor, der Erscheinungsort, der Verlag und weitere Angaben.

Der Umfang beider Dateien war in Bezug auf die Übermittlung und weitere Verarbeitung vorteilhaft. Als gepackte Files waren sie mit ca. 1,3 MB bzw. 1,9 MB leicht als Anhang per E-Mail zu verschicken und konnten mit dieser Größe auch

beim Parsen mit einer durchschnittlichen bis guten Rechnerleistung in einer akzeptablen Zeit bewältigt werden.

Grundsätzlich muss es als glücklicher Umstand eingeschätzt werden, dass man auf so einfachem Weg an den gesamten Bestand des kid gelangt ist. Nach Angaben von Herrn Pretz (siehe Gesprächsprotokoll Pretz v. 04.12.2001, Anlage 20) scheuen sich viele Bibliothekare, Gesamtbestände oder auch nur Teilbestände aus Themengebieten in Form einer Textdatei verfügbar zu machen. Offenbar besteht hier die Tendenz, selber erschlossene Dokumentenbestände in ihrer Gesamtheit nicht von Dritten einsehen zu lassen. Aus diesem Grund ist das Verhalten von Herrn Schmitz-Veltin als sehr kooperativ zu bewerten.

Der beabsichtigte automatische Abgleich der Datenbestände sollte nun in folgenden Schritten durchgeführt werden:

- Parsen sowohl der Dateien mit Beständen des kid als auch der mit den relevanten Hildesheimer Beständen, so dass daraus zwei in ihrer Struktur gleiche Textdateien resultieren
- Vergleich der Inhalte der beiden gewonnen Dateien auf Identität und Ausgabe der dabei gefundenen Hildesheimer Titel

### **4.3.1 Das Parsen der Literaturbestände**

#### *4.3.1.1 Die Programmiersprache gawk*

Auf der Suche nach einem geeigneten Tool zur Bearbeitung der vorliegenden Dateien (Bestand des kid und Titel der UB Hildesheim) stellte sich die Notwendigkeit, eigene Textparser zu schreiben, die die anstehenden Aufgaben erfüllen konnten. Dafür wurde die Programmiersprache gawk herangezogen, die im Rahmen des GNU-Projektes als freie Software angeboten wird.

Das GNU-Projekt (GNU ist eine rekursive Abkürzung von „GNU's Not Unix!“) wurde 1984 begonnen, um ein vollständiges Unix-artiges Betriebssystem zu entwickeln, das freie Software zur Verfügung stellt (vgl. Arnold, 2002). Es soll der Zusammenarbeit der „computing community“ dienen, und Schwierigkeiten

bei der Weiterentwicklung von Software überwinden helfen, die durch Eigentumsrechte an Software-Produkten entstanden sind (vgl. Stallman, 2001). Das GNU-Projekt geht zurück auf eine Initiative der Free Software Foundation (FSF), deren Gründer Richard M. Stallman ist (vgl. Robbins 2001, The GNU Project and This Book).

Der ursprüngliche Name der Programmiersprache awk stammt von den Initialen ihrer Entwickler: Alfred V. Aho, Peter J. Weinberger und Brian W. Kernighan. Die erste Version wurde 1977 in den AT&T Bell Laboratories geschrieben. Eine 1985 veröffentlichte neue, leistungsfähigere Version fand über das Unix System V Release 3.1 (SVR3.1) weite Verbreitung. 1986 wurde awk in das GNU-System implementiert und erhielt in dieser Version die Bezeichnung gawk, welche für „GNU awk“ steht (ganz korrekt: `@command{gawk}`). Die Sprache wird von vielen UNIX-Systemen unterstützt und läuft auf einer Vielzahl von anderen Betriebssystemen: Mac OS X, MS-DOS, Microsoft Windows (alle Versionen) und OS/2 PC's, Atari und Amiga micro-computers, BeOS, Tandem D20, und VMS (vgl. Robbins 2001, History of `@command{awk}` and `@command{gawk}`). Im Rahmen dieser Arbeit wurde das Betriebssystem Windows98 verwendet.

Gawk wurde für Benutzer geschrieben, die häufig mit verschiedensten Textdateien arbeiten. Es hilft dabei, gezielte, sich wiederholende Änderungen in gleichen Textteilen vorzunehmen oder Daten aus Teilen von ausgesuchten Zeilen zu entnehmen. Im Gegensatz zu einem gleichwertigen Programm in C, C++ oder Pascal, dessen Erstellung mit einem hohen Arbeitsaufwand verbunden wäre, reichen bei gawk dafür wenige Programmzeilen aus. Im Einzelnen kann gawk bei folgenden Aufgaben hilfreich sein:

- Verwaltung von kleineren Datenbeständen
- Erzeugung von Berichten
- Beurteilung von Daten
- Erstellung von Indexen und anderen dokumentvorbereitenden Arbeiten
- Experimentieren mit Algorithmen, die man später an andere Programmiersprachen anpassen kann
- Extrahieren von Daten für eine weitere Verarbeitung

(vgl. Robbins 2001, Preface)

Des Weiteren wurde für die Bearbeitung der Textdateien eine spezielle Software in einer Testlizenz verwendet, der Texteditor UltraEdit der Firma IDM Computer Solutions, Inc. in der Version UltraEdit-32 - v7.10. Er war insbesondere deshalb für die Durchführung der Arbeiten sehr wichtig, da er unbegrenzte Dateigrößen unterstützt. Außerdem bietet er die Möglichkeit, Dateien auf Dubletten hin zu prüfen und sie zu sortieren. Auch die Inhalte zweier Dateien können verglichen werden, die Suchen-Ersetzen-Funktion kann über mehrere Dateien hinweg durchgeführt werden und die Textumwandlung in Groß- oder Kleinbuchstaben kann problemlos für große Dateien erfolgen. Durch seine Syntaxhervorhebung ist der UltraEdit auch als HTML-Editor geeignet.

#### 4.3.1.2 Parsen und Aufbereiten des kid

Der erste Schritt bestand in einem Parsen der beiden Dateien des Konstanzer kid-Bestandes. Die Zeilen des ersten Files wiesen die folgende Struktur auf:

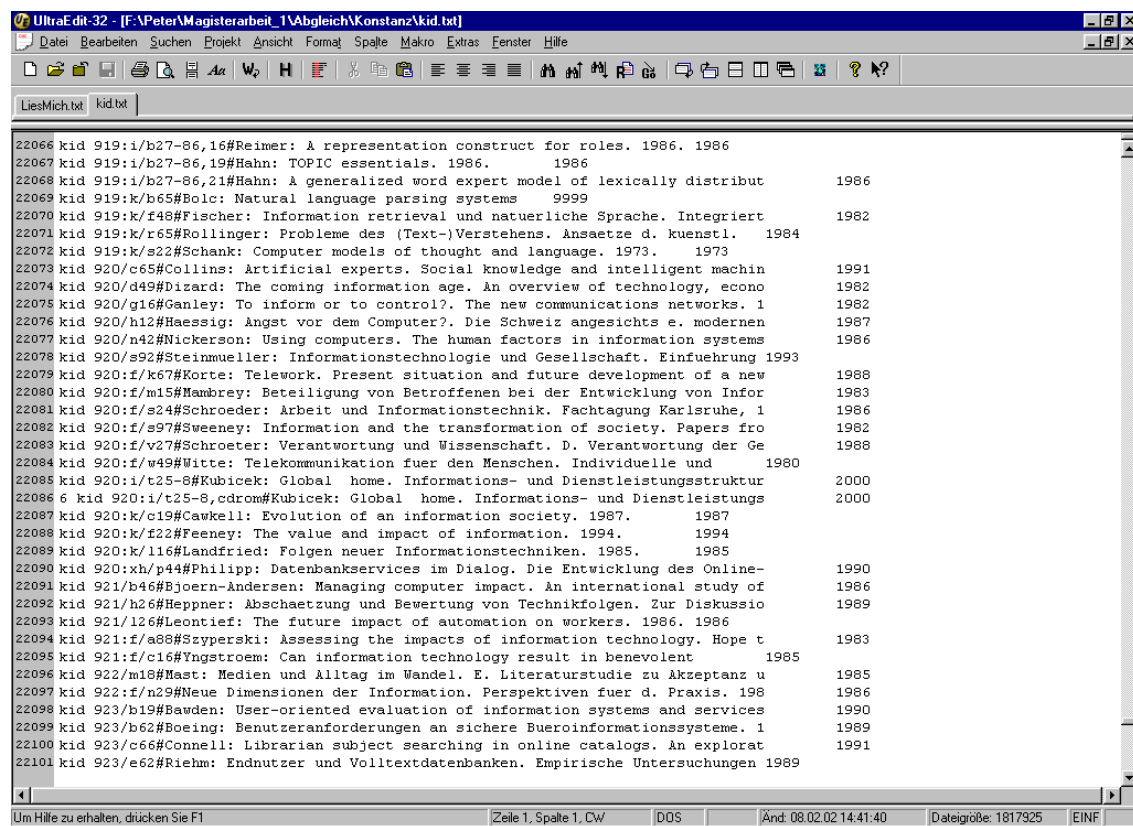


Abbildung 8: Auszug aus der ersten Konstanzer Bestandsdatei (Titel mit kid Signaturen)



Der erste Eintrag stellt die Signatur dar, hinter dem „#“ folgt der Autor, anschließend kommt der vollständige oder abgebrochene Titel (je nach Länge), gefolgt vom Erscheinungsjahr und dem Jahr der Anschaffung. Alle Informationen zu einem Titel finden sich in derselben Zeile. Aus diesen Daten galt es, die relevanten heraus zu filtern.

Zuerst einmal mussten die Signaturen entfernt werden, da sie für den späteren Abgleich unwichtig war. Dafür wurde im Parser `pars_k1.awk` (siehe Anlage 6) das Zeichen „#“ zum neuen Separator zwischen den sogenannten „Records“ erklärt. Per Default-Einstellung ist der Separator der Zeilenumbruch, so dass jede einzelne Zeile ein eigener Record ist (vgl. Robbins 2001, How Input Is Split into Records). Durch die Änderung begann jetzt nach jedem „#“ eine neue Zeile. Das Ergebnis wurde in eine Textdatei geschrieben.

Anschließend wurde im Parser `pars_k2.awk` (siehe Anlage 7) diese neue Datei dahingehend verändert, dass die Zeilen mit der Signatur eliminiert wurden. Übrig blieben also nur der Autor, der Titel und eine bzw. zwei Jahreszahlen. Letztere mussten ebenfalls noch aus den Zeilen entfernt werden, da sie in der Zieldatei (Hildesheimer Bestand) nicht auftauchten und so eine Identitätsprüfung fehl geschlagen wäre. Dadurch gingen zwar die mit einer Jahreszahl beginnenden Titel verloren, was aber aufgrund des seltenen Vorkommens nur wenige Dokumente betrifft (siehe Parser `pars_k2.awk`, Anlage 7). Auch die gesamte Zeilenlänge musste auf die Zieldatei angepasst werden (55 Zeichen) (siehe Parser `pars_k3.awk`, Anlage 8); sie sollte deren Zeilenlänge nicht überschreiten, da beim Abgleich unter anderem geprüft werden sollte, ob die Zeile des Konstanzer kid Bestandteil der Zeile des Hildesheimer Bestandes ist.

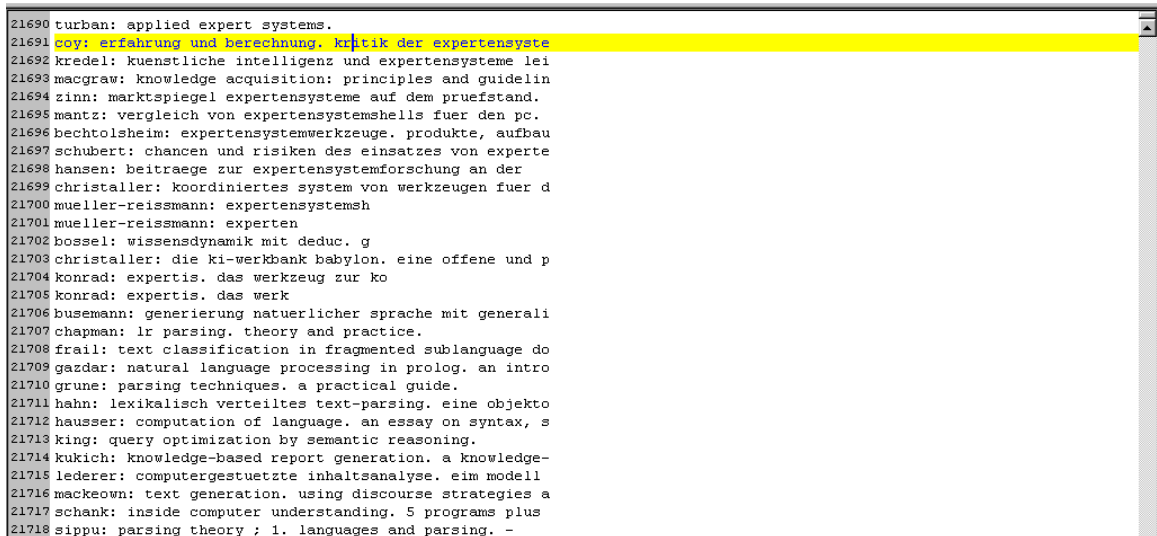
Bei den ersten Testdurchläufen des Abgleichs (mit 1.000 Konstanzer Titeln und der gesamten Hildesheimer Bestandsdatei) wurden neue Probleme deutlich, die vorher nicht erkannt worden waren und im Nachhinein, vor dem endgültigen Abgleich, beseitigt werden mussten:

- Titel, die aufgrund der abgebrochenen Form nur eine aufgehende Klammer enthielten, ergaben beim Parsen eine Fehlermeldung. Offenbar können sie von den weiter unten erläuterten „regular expressions“ (siehe Kapitel 4.3.2)

nicht interpretiert werden. Deshalb wurden alle Titelbestandteile hinter einer sich öffnenden Klammer entfernt, wenn diese in der betroffenen Zeile nicht wieder geschlossen wurde (siehe Parser pars\_k3.awk, Anlage 8).

- Bei einigen Zeilen fand sich am Ende das Tabulatorzeichen „>“, welches direkt ohne Leerschritt auf das letzte vollständige oder abgebrochene Wort folgte. Die für den Abgleich verwendeten „regular expressions“ interpretieren offensichtlich dieses Formatierungszeichen und verhindern dadurch das Auffinden in der Hildesheimer Bestandsdatei. Um hier Abhilfe zu schaffen, wurden über die Funktion „Suchen - Ersetzen“ des verwendeten Texteditors alle Tabulatorzeichen in der Datei entfernt.
- Einige Leerzeilen mussten auf die gleiche Weise wie die Tabulatorzeichen entfernt werden, da sie beim Parsen eine Fehlermeldung ergaben.
- Als Schwierigkeit erwiesen sich ebenfalls Abweichungen bei der Groß- und Kleinschreibung am Wortanfang, beispielsweise bei dem Wort „Internet“. Aus diesem Grund wurden alle Buchstaben des Textes über die entsprechende Funktionalität des Editors in Kleinbuchstaben umgewandelt (genau wie bei der Hildesheimer Bestandsdatei).
- Stichpunktartige Untersuchungen des geparsten kid ließen keine weiteren Probleme erkennen.

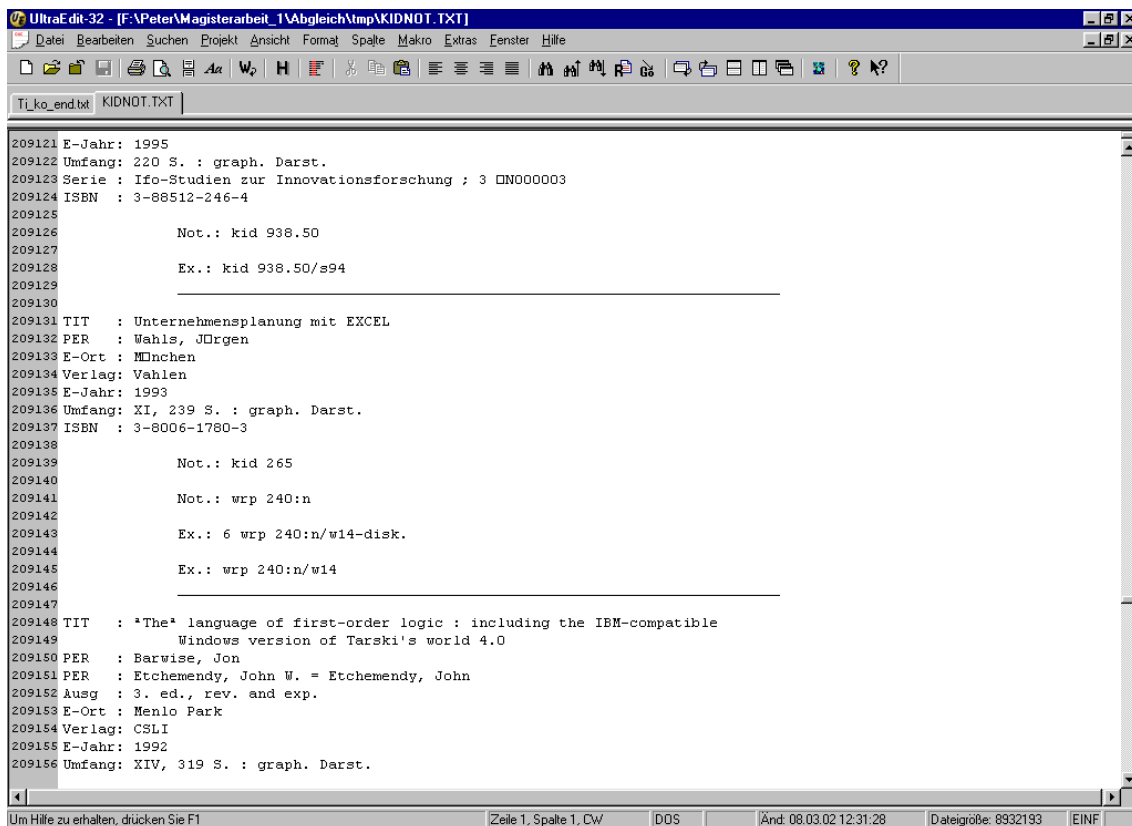
Die Ergebnisdatei hatte jetzt folgenden Inhalt:



```
21690 turban: applied expert systems.
21691 soy: erfahrung und berechnung. kritik der expertensysteme
21692 kredel: kuenstliche intelligenz und expertensysteme lei
21693 macgraw: knowledge acquisition: principles and guidelin
21694 zinn: marktspiegel expertensysteme auf dem pruefstand.
21695 mantz: vergleich von expertensystemshells fuer den pc.
21696 bechtolsheim: expertensystemwerkzeuge. produkte, aufbau
21697 schubert: chancen und risiken des einsatzes von experte
21698 hansen: beitraege zur expertensystemforschung an der
21699 christaller: koordiniertes system von werkzeugen fuer d
21700 mueller-reissmann: expertensystemsh
21701 mueller-reissmann: experten
21702 bossel: wissensdynamik mit deduc. g
21703 christaller: die ki-werkbank babylon. eine offene und p
21704 konrad: expertis. das werkzeug zur ko
21705 konrad: expertis. das werk
21706 busemann: generierung natuerlicher sprache mit generali
21707 chapman: lr parsing. theory and practice.
21708 frail: text classification in fragmented sublanguage do
21709 gazdar: natural language processing in prolog. an intro
21710 grune: parsing techniques. a practical guide.
21711 hahn: lexikalisch verteiltes text-parsing. eine objekto
21712 hausser: computation of language. an essay on syntax, s
21713 king: query optimization by semantic reasoning.
21714 kukich: knowledge-based report generation. a knowledge-
21715 lederer: computergestuetzte inhaltsanalyse. ein modell
21716 mackeown: text generation. using discourse strategies a
21717 schank: inside computer understanding. 5 programs plus
21718 sippu: parsing theory ; 1. languages and parsing. -
```

Abbildung 9: Auszug aus der geparsten Konstanzer Bestandsdatei des kid (Signaturen)

Die zweite Konstanzer Datei, die neben den kid-Signaturen auch die Titel mit einer zugeordneten kid-Notation (in Konstanz als Systemstelle bezeichnet) enthielt, sah folgendermaßen aus:



```

UltraEdit-32 - [F:\Peter\Magisterarbeit_1\Abgleich\mp\KIDNOT.TXT]
Datei Bearbeiten Suchen Projekt Ansicht Format Spalte Makro Extras Fenster Hilfe

Ti_ko_end.txt KIDNOT.TXT

209121 E-Jahr: 1995
209122 Umfang: 220 S. : graph. Darst.
209123 Serie : Ifo-Studien zur Innovationsforschung ; 3 □N0000003
209124 ISBN : 3-88512-246-4
209125
209126 Not.: kid 938.50
209127
209128 Ex.: kid 938.50/s94
209129
209130
209131 TIT : Unternehmensplanung mit EXCEL
209132 PER : Wahls, Jürgen
209133 E-Ort : München
209134 Verlag: Vahlen
209135 E-Jahr: 1993
209136 Umfang: XI, 239 S. : graph. Darst.
209137 ISBN : 3-8006-1780-3
209138
209139 Not.: kid 265
209140
209141 Not.: wrp 240:n
209142
209143 Ex.: 6 wrp 240:n/w14-disk.
209144
209145 Ex.: wrp 240:n/w14
209146
209147
209148 TIT : *The* language of first-order logic : including the IBM-compatible
209149 Windows version of Tarski's world 4.0
209150 PER : Barwise, Jon
209151 PER : Etchemendy, John W. = Etchemendy, John
209152 Ausg : 3. ed., rev. and exp.
209153 E-Ort : Menlo Park
209154 Verlag: CSLI
209155 E-Jahr: 1992
209156 Umfang: XIV, 319 S. : graph. Darst.
  
```

Abbildung 10: Auszug aus der zweiten Konstanzer Bestandsdatei (Titel mit kid-Signaturen („Ex.“) und kid-Systemstelle („Not.“))

Die erste Aufgabe eines Parsers bestand hier darin, den jeweiligen Titel und Autor eines Datensatzes auszulesen. Für diesen Zweck wurden die Einträge der Zeilen „TIT“ und „PER“ nach dem Doppelpunkt herausgeparst und formatiert in eine Datei ausgegeben (siehe Parser pars\_k4.awk, Anlage 9). Dabei musste beachtet werden, dass bei Mehrfachnennungen von Autoren nur der erste Nachname Eingang in die Ergebnisdatei fand, was durch den Einbau der Variable „es\_gibt\_einen\_verfasser“ gelang. Dass für den Abgleich nur dieser erste Nachname relevant war, hatte sich zwischenzeitlich beim Parsen des Hildesheimer Bestandes ergeben und wird an dortiger Stelle genauer erläutert und belegt (siehe Kapitel 4.3.1.3).

Die aus diesem ersten Arbeitsschritt resultierende Datei wurde durch die beiden folgenden Parser (siehe pars\_k5.awk und pars\_k6.awk, Anlage 10 und 11)

dahingehend verändert, dass die Zeilenlänge auf 55 Zeichen begrenzt wurde und dass der Zeileninhalt nach aufgehenden runden (siehe Parser `pars_k5.awk`, Anlage 10) und eckigen Klammern (siehe Parser `pars_k6.awk`, Anlage 11), die in der betreffenden Zeile nicht wieder geschlossen wurden, entfernt wurde. Diese Problematik ist bereits beim Parsen der ersten Konstanzer Datei beschrieben worden. Außerdem erfolgte, wie auch in der ersten kid-Datei, eine Umstellung des kompletten Zeicheninhalt auf Kleinschreibung und einige Leerzeilen mussten per Hand entfernt werden.

Im Unterschied zum ersten kid-File wurden noch weitere Korrekturen vorgenommen. So zeigte sich, dass die zweite Datei einige Sonderzeichen (wie z.B. <sup>a</sup>, š, □), enthielt, die entsprechend umgewandelt bzw. entfernt werden mussten. Die meisten dieser Zeichen standen stellvertretend für Umlaute, die, analog zur Bearbeitung der ersten kid-Datei, in der Form „ue“, „ae“ und „oe“ über die entsprechende Funktion des Editors eingefügt wurden. Auch fanden sich in einigen Titeln Untertitel, die durch einen Doppelpunkt eingeleitet wurden, vor dem sich ein Leerschritt befand. Da in der ersten Konstanzer Bestandsdatei dieser Leerschritt nicht auftauchte, wurde er in der gesamten Datei durch den Editor entfernt.

Die zweite kid-Bestandsdatei hatte nun die gleiche Struktur wie die erste und konnte an diese angehängt werden. Die neue Datei, welche nun für den Abgleich bereit stand, umfasste insgesamt 42.149 Zeilen, wobei darin viele Titel doppelt enthalten waren (ohne dass sie eine völlig identische Zeichenfolge gehabt hätten, so dass man sie nicht ohne größeren Aufwand hätte entfernen können).

##### *4.3.1.3 Parsen und Aufbereiten des Hildesheimer Bestandes*

Die, wie im Kapitel 4.2 beschrieben, zusammen gestellte Textdatei mit den informationswissenschaftlichen Titeln der Universitätsbibliothek Hildesheim war im Vergleich zur geparsten Datei des kid weniger homogen. Es fanden sich dort eine Vielzahl von nicht relevanten Daten. Auch war die Datei mit ihren ca. 13 MB wesentlich größer. Ein Textausschnitt verdeutlicht die vorliegende Struktur:

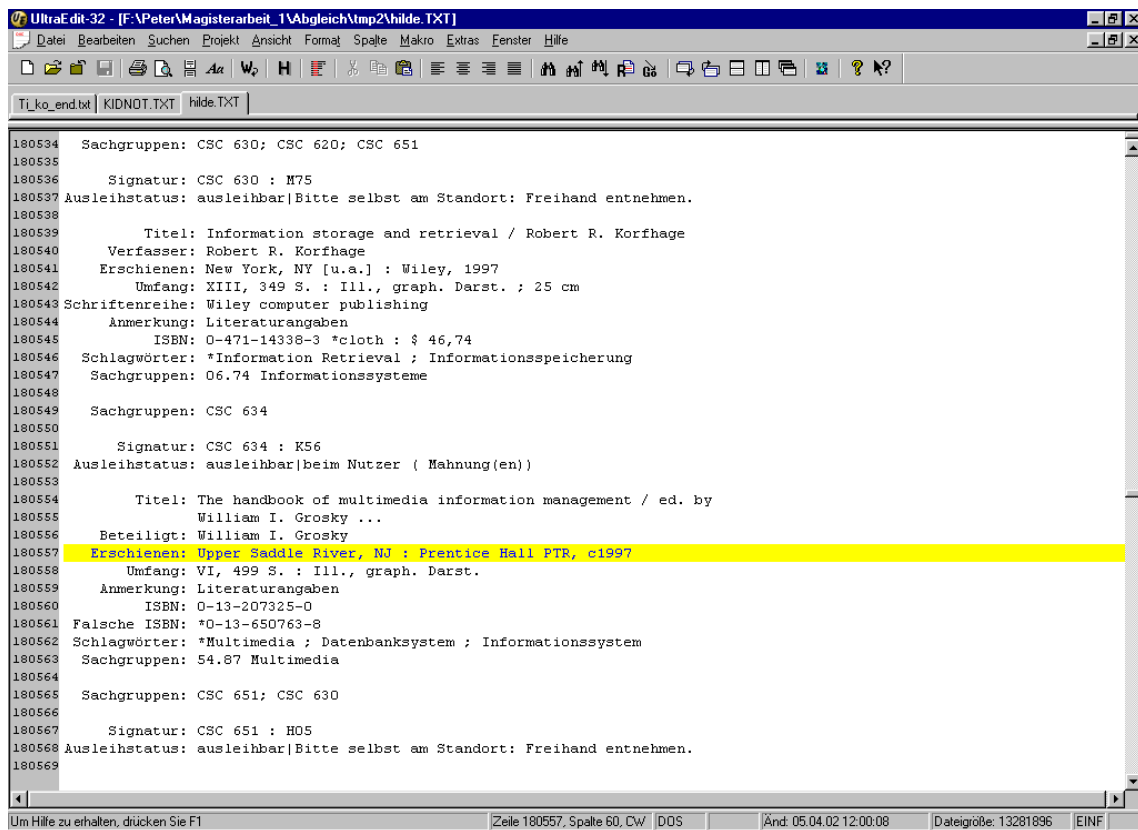


Abbildung 11: Auszug aus der Hildesheimer Bestandsdatei

Die denselben Titel betreffenden Informationen beginnen jeweils mit der Zeile „Titel:“ und enden mit „Ausleihstatus:“. Die sich dazwischen befindenden Daten sollten für den Abgleich in einer adäquaten Form aufbereitet werden. Dabei musste bedacht werden, dass anstatt „Titel:“ häufig auch die Bezeichnung „Mehrteiliges Werk:“ benutzt wird, an der Stelle von „Verfasser:“ findet man oft „Beteiligt:“.

Damit die gleiche Struktur wie die der Ergebnisdatei des geparsten kid entstand, mussten die Zeilen mit folgender einleitender Bezeichnung aus der Datei herausgelesen werden: „Titel:“ bzw. „Mehrteiliges Werk:“ und „Beteiligt:“ bzw. „Verfasser:“. Bei Letzteren musste beachtet werden, dass bei nicht wenigen Titeln beide Bezeichnungen auftauchten. Da in diesen Fällen nur der Verfasser durch den Parser aufgefunden werden sollte, musste gewährleistet sein, dass die Einträge hinter „Beteiligt:“ keinen Eingang in die Ergebnisdatei fanden. Dabei wurde wiederum die Variable „es\_gibt\_einen\_verfasser“ verwendet (siehe Parser pars\_h08.awk, Anlage 12). Eine weitere Problematik bei der Erfassung des Autors war, dass hinter einigen Verfassern eine Jahresangabe stand

(z.B. \*1952-\*), vermutlich eine Angabe zum Geburtsjahr des Autors. Hier musste darauf geachtet werden, dass auch wirklich der Nachname der Person und nicht die Jahresangabe, wie anfänglich geschehen, vom Parser gefunden wurde.

Außerdem sollten, zwecks der Erleichterung bei einer späteren Zuordnung der gefundenen Hildesheimer Titel auf die neue Klassifikation, die Zeilen „Sachgruppen:“ und „Signatur:“ mit in die Datei für den Abgleich aufgenommen werden. Die Bezeichnung „Sachgruppen:“ taucht häufig zweimal pro Titel auf: Bei der ersten Erwähnung werden Kategorien der Basisklassifikation aufgeführt, unter denen der Titel virtuell ebenfalls zu finden ist. Bei der zweiten Erwähnung trifft man auf Hildesheimer Notationen, unter denen der darüber beschriebene Titel, ebenfalls digital zu finden ist. Es handelt sich also auch hier durch die Eintragung dieser sogenannten Nebenstellen um eine Art virtuelle Zuweisung. Unter diesen Notationen kann sich, wie bereits erwähnt, auch die der Signatur befinden. Diese findet man in voller Länge unter „Signatur:“. Neben diesen vollständigen Angaben finden sich auch die sogenannten Minimalsätze (siehe Kapitel 4.2), welche nur aus Signatur und Ausleihstatus bestehen. Sie sind für den Abgleich ohne Bedeutung.

In der Zeile „Titel:“ sind außer dem vollen Titel des Buches, wie in Abbildung 11 zu sehen, auch der oder die Autoren zu finden. Wie im obigen Beispiel, können sich die diesbezüglichen Informationen über mehrere Zeilen erstrecken. Da aber für den Abgleich eine kurze Titelzeile ausreicht (die Datei des kid enthält auch nur abgebrochene Titel), wurde beim Parsen nur die erste Zeile berücksichtigt (siehe Parser pars\_h08.awk, Anlage 12).

Beim Auffinden der Autoren in der Zeile „Verfasser:“ bzw. „Beteiligt:“ ergab sich das Problem, dass es nicht nur einen sondern auch mehrere Autorennennungen gab. In jedem Fall durfte beim Parsen nur der Nachname Berücksichtigung finden. Um herauszufinden, nach welchem System beim kid zwischen mehreren Autoren ausgewählt wurde, wurden stichpunktartige Untersuchungen durchgeführt. Wie die fünf ausgewählten Beispiele unten belegen, wird im kid immer nur der erst genannte Autor eines Werkes aufgeführt (siehe

Abbildung 12). Aus diesem Grund wurde auch nur der jeweils erste der Autoren mit seinem Nachnamen in die Datei für den Abgleich aufgenommen (der Name, der bei mehreren Autoren vor dem Semikolon steht bzw. bei nur einem Autor der letzte Name der Zeile).

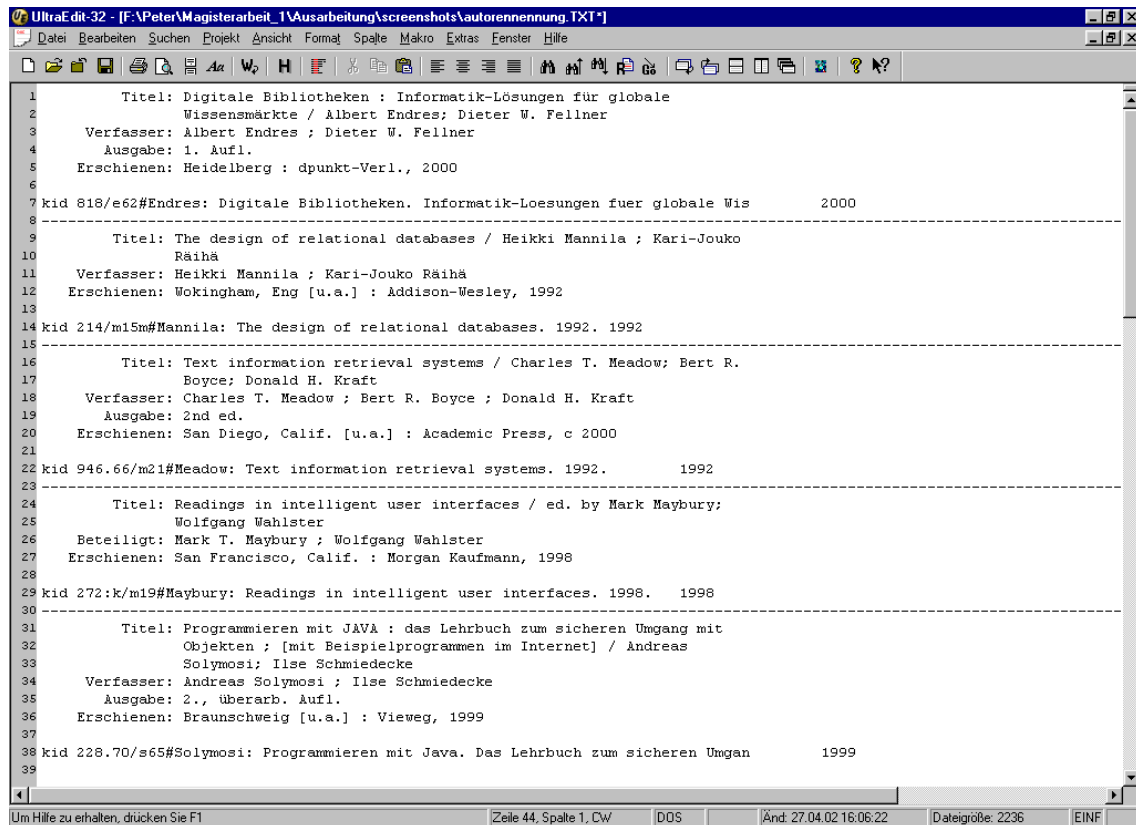


Abbildung 12: Fünf Beispiele für die Verwendung des jeweils erst genannten Autors der Hildesheimer Bestandsdatei im kid  
(Der jeweils obere Titel stammt aus der Hildesheimer Bestandsdatei, die jeweils unterste Zeile aus dem kid.)

Wie schon erwähnt, findet sich die Bezeichnung „Sachgruppen:“ bei vielen Titeln zweimal. Da nur die Zeile mit den Hildesheimer Notationen von Belang für eine Zuordnung der gefundenen Titel in die neue Klassifikation war, wurde auch nur diese beim Parsen berücksichtigt. Die Signatur der Titel wurde in voller Länge erfasst.

Wie schon beim kid stellten sich nach ersten Testdurchläufen des Abgleichs bis dahin unbekannte Probleme heraus, die behoben werden mussten:

- Tauchte im Titel das Zeichen „:“ auf, so stand es nicht direkt im Anschluss an das vorhergehende Wort, sondern mit einem Leerschritt Abstand und wurde somit als eigenes Feld interpretiert. Dadurch wurde dieser Titel beim Abgleich nicht gefunden. Zur Lösung des Problems wurde, wiederum mit der Suchen - Ersetzen - Funktion des Texteditors, jeweils der Leerschritt vor dem Doppelpunkt entfernt.
- Als problematisch stellte sich ebenso die unterschiedliche Verwendung von Umlauten in den beiden abzugleichenden Dateien heraus. Im Gegensatz zu den Hildesheimer Titeln werden im kid „ä“, „ü“ und „ö“ als „ae“, „ue“ und „oe“ dargestellt. Um für Einheitlichkeit zu sorgen, wurden in der Hildesheimer Bestandsdatei alle Umlaute entsprechend umgewandelt (wiederum mit der Suchen - Ersetzen - Funktion). Die Anpassung wurde nicht im kid durchgeführt, um zu vermeiden, dass ungewollte Zusammenziehungen von Buchstaben erfolgen, z.B. bei Wörtern wie microelectronic.
- Um die aus der unterschiedlichen Groß- und Kleinschreibung resultierenden Probleme zu beseitigen, wurden alle Buchstaben der Hildesheimer Titel in Kleinbuchstaben umgewandelt (wie schon zuvor im kid).
- Auch hier ließen stichpunktartige Untersuchungen der geparsten Hildesheimer Datei keine zusätzlichen Probleme erkennen.

Am Ende wurden aus Gründen der Übersichtlichkeit noch mit dem Editor die Titel in aufsteigender Folge alphabetisch sortiert und die Dubletten entfernt. Ein Ausschnitt zeigt die neue Struktur der geparsten Daten:



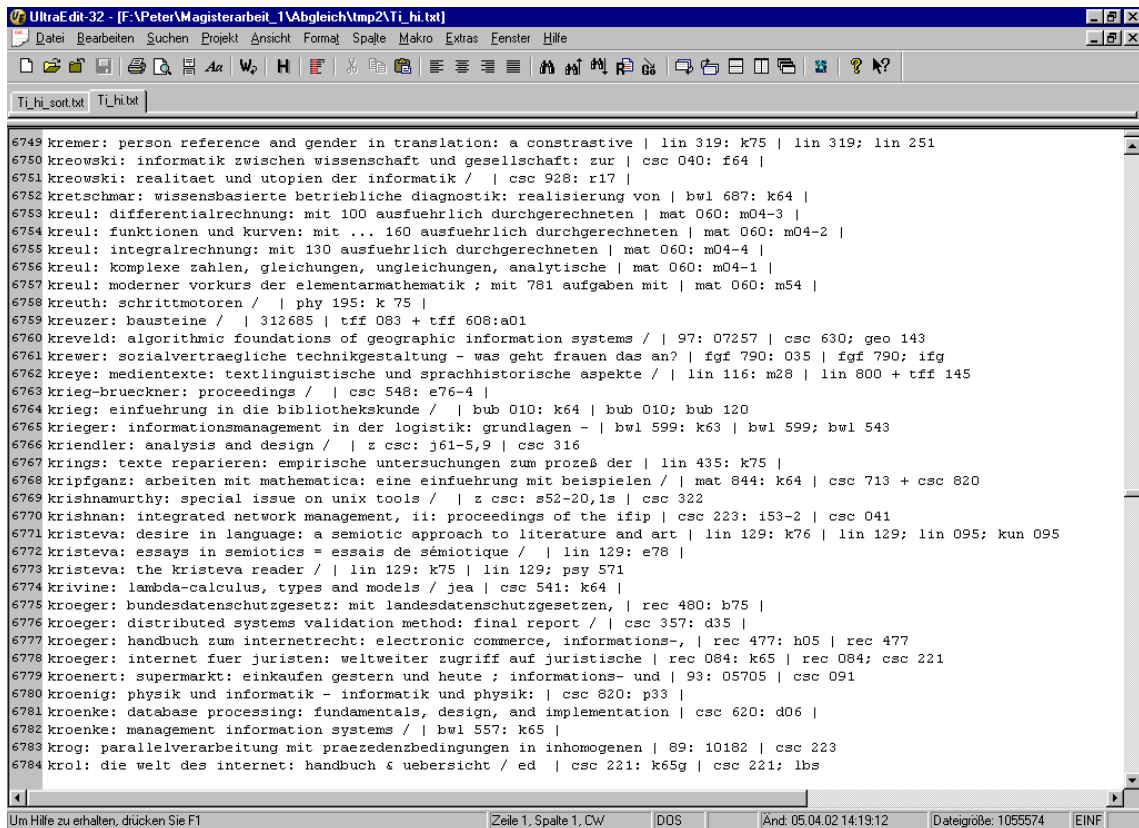


Abbildung 13: Ausschnitt aus der bearbeiteten Hildesheimer Bestandsdatei  
(Autor: Kurztitel | Signatur | Nebenstellen)

### 4.3.2 Der Abgleich des Konstanzer und Hildesheimer Bestandes

Bei dem nun folgenden Abgleich der beiden zuvor bearbeiteten Dateien sollte geprüft werden, ob sich in ihnen identische Titel befinden. Zu diesem Zweck sollte die Konstanzer Datei zeilenweise ausgelesen und mit den vorliegenden Hildesheimer Titeln verglichen werden. Die dabei gefunden Titel sollten gesondert in eine Datei ausgegeben und dort weiter bearbeitet werden.

Um dies zu erreichen, wurde im ersten Teil des Parsers (siehe pars\_kh9.awk, Anlage 13) in einer While-Schleife über das Kommando „getline“ die erste Zeile aus der Konstanzer Datei eingelesen. Im nächsten Schritt wurde den ersten vier Feldern dieser Zeile (\$1 bis \$4) – ein Feld entspricht den Zeichen zwischen zwei Leerschritten, \$0 steht in gawk für eine gesamte Zeile – eine Variable zugewiesen; davon repräsentiert das erste Feld den Nachnamen des Autors und die restlichen drei die ersten Worte des Titels.

In der nun folgenden zweiten While-Schleife wird die erste Zeile aus der bearbeiteten Hildesheimer Bestandsdatei gelesen und analog zur Konstanzer Zeile erfolgt eine Belegung der ersten vier Felder durch Variablen. Für diese Hildesheimer Zeile wird dann geprüft, ob sie ein hohes Maß an Ähnlichkeit mit der zuvor eingelesenen Konstanzer Zeile aufweist. Ist das der Fall, so wird die Zeile in eine Datei geschrieben und der Parser „holt“ sich die nächste Hildesheimer Zeile. Er tut dies so lange, bis die komplette Hildesheimer Datei durchlaufen ist. Dann liest er die nächste Konstanzer Zeile ein und der Ablauf wiederholt sich. Nach Prüfung der letzten Konstanzer Zeile ist der Vorgang beendet.

Der eigentliche Abgleich geschieht in der zweiten While-Schleife, indem, bezogen auf den Inhalt der vorliegenden Zeilen, verschiedene Konditionen geprüft werden. Gawk bietet die Möglichkeit, über sogenannte „regular expressions“ Zeilen auf identische Inhalte hin zu untersuchen. Eine solche „regular expression“ beschreibt eine Folge von Strings bzw. Feldern. Man kann dafür entweder eine Konstante verwenden, so dass immer nach der gleichen Folge von Zeichen gesucht wird, oder aber auch eine Variable, eine sogenannte „dynamic regular expression“. Letzteres geschieht beim Abgleich der beiden Dateien, da jeweils die zu prüfende Konstanzer Zeile als Variable eingesetzt wird. In der ersten Bedingung wird also untersucht, ob die Inhalte der Konstanzer Zeile sich in der Hildesheimer Zeile wiederfinden, Teile der beiden Zeilen somit identisch sind.

Allerdings ist diese Art des Vergleichs nicht ausreichend. Auf diese Weise werden Titel, die bis auf ein oder mehrere Worte identisch sind, nicht erfasst. Da solche Abweichungen auch dann auftauchen, wenn es sich um das gleiche Buch handelt, mussten weitere Konditionen für den Abgleich angelegt werden. Ansonsten würden z.B. Titel, bei denen die Konstanzer Zeile länger als die Hildesheimer Zeile ist (weil sie zusätzliche Informationen beinhaltet), nicht erfasst werden. Genauso kam es vor, dass einige Konstanzer Titel am Ende der Zeile andere Informationen (z.B. Jahreszahlen oder Orte) enthielten, als die ihnen entsprechenden Hildesheimer Titel. Auch Büchertitel, bei denen in einer der beiden Dateien eine Abkürzung vorgenommen wurde, fielen so beim Abgleich nicht auf. Gleiches gilt für die sinnlose Buchstabenkombinationen „ieee“, die

stellvertretend für ein anderes Wort steht und sich öfter in der Konstanzer Datei findet. Als Vorteil erweist sich hingegen, dass Sonderzeichen wie, z.B. „\*“, „?“ und „.“, von den „regular expressions“ nicht interpretiert werden, sondern in diesen Fällen nur auf identische Wortlänge geprüft wird. Wenn sich in dem entsprechenden Wort der zu prüfenden Zeile statt eines Doppelpunktes ein einfacher Punkt befindet (was häufiger vorkommt), so wird das Wort trotzdem als identisch erkannt.

Um die aufgezählten Nachteile der Suche mit „regular expressions“ auszuschalten, wurde bei den weiteren Bedingungen untersucht, ob mindestens drei von vier Wörtern bzw. Feldern des Titels identisch sind. Auf diese Weise werden unter Umständen auch solche Titel erfasst, die sich sehr ähnlich sind, ohne dass es sich zwangsläufig um den gleichen Titel handelt. Die Testdurchläufe mit etwa 1000 Konstanzer Titel ergaben allerdings, dass dabei in der Regel nur informationswissenschaftliche Bücher als Treffer geliefert wurden, was insbesondere auf den jeweils identischen Autor zurückzuführen ist. Solche Titel wären ohne diese Art von Kondition sonst gar nicht erfasst worden.

Der Nachteil einer vorgegebenen Identität von mindestens drei Feldern besteht darin, dass man sehr kurze Zeilen, die nur aus dem Autorennamen und einem Wort im Titel bestehen, so nicht finden würde. Die in diesem Fall nicht belegten Felder verhindern eine Erfüllung der Kondition. Glücklicherweise wurden diese kurzen Titel aber über den Abgleich mit den „regular expressions“ ausfindig gemacht.

Auch Titel der Konstanzer Datei, bei denen die Autorenangabe fehlte, wurden durch diesen Parser ausfindig gemacht. Die letzte Kondition der If-Schleife verlangte allerdings, dass mindestens vier Felder der verglichenen Zeilen identisch sein müssen. Versuche mit einer geringeren Anzahl von Feldern hatten ergeben, dass sich dabei aufgrund des fehlenden Autors eine hohe Trefferquote ergab, die sehr viele nicht relevante, d.h. nicht informationswissenschaftliche Bücher, enthielt.

Aufgrund der Größe der beiden abzugleichenden Files (mehr als 42.000 bzw. 12.000 Zeilen) wurde für die Durchführung ein Rechner mit der entsprechenden Performance ausgewählt. Ein in der Universität Hildesheim aufgestellter Rechner verfügte mit einem Hauptspeicher von einem GB und zwei Prozessoren mit jeweils 800 MHz über die notwendigen Ressourcen. Die gesamte Rechenzeit für den Durchlauf des Programms betrug in etwa vier Stunden.

Aus der Ergebnisdatei wurden die Dubletten, die sich aus Mehrfachnennungen von Titeln in der Konstanzer Datei ergaben, entfernt und eine Sortierung in aufsteigender Folge vorgenommen. Danach enthielt die Datei 4.905 gefundene Hildesheimer Titel.

#### **4.4 Literaturanschaffungen im Institut für Angewandte Sprachwissenschaft**

Von Anfang an bekannt war die Problematik, dass durch den Abgleich mit dem Bestand einer anderen Bibliothek nicht alle informationswissenschaftlichen Bücher in der UB Hildesheim erfasst werden konnten. Auch die Literaturrecherche hatte gezeigt, dass insbesondere einige Bücher mit thematischem Schwerpunkt auf Lehrgebieten der Hildesheimer Informationswissenschaft nicht in Konstanz zu finden waren.

Um auch möglichst viele von diesen Büchern mit in die Klassifikation aufnehmen zu können, wurden die Buchbestellungen mit herangezogen, die von Dozenten mit informationswissenschaftlichen Lehrveranstaltungen in den letzten Jahren (seit 1998) aufgegeben wurden. Bei den Dozenten handelt es sich um Prof. Dr. Christa Womser-Hacker und Dr. Thomas Mandl. Dr. Folker Caroli steuerte noch eine Liste von Bibliotheksbüchern aus den Handapparaten (Stand: 14.8.2001) bei.

Mit den dort aufgeführten Titeln wurden Anfragen an den OPC4 gestartet und die gefundenen Dokumente wurden in einer gemeinsamen Datei abgespeichert. Diese wurde genauso wie die Hildesheimer Bestandsdatei bearbeitet, um eine

Einheitlichkeit mit den Titeln des Abgleichs zu herbeizuführen: Sie wurde geparst, die Umlaute wurden entfernt und die Schreibweise wurde auf Kleinbuchstaben umgestellt. Anschließend wurden die Titel zu der Ergebnisdatei des Abgleichs hinzugefügt.

### **4.5 Suchanfragen an den Web-OPAC der UB Hildesheim**

Nach Zuweisung der beim Abgleich ermittelten Hildesheimer Titel auf die neue Klassifikation (siehe Kapitel 5.2) hatte sich gezeigt, dass nicht alle Themengebiete mit Literatur versorgt werden konnten. Einige von Ihnen haben keine Entsprechung in der Beschreibung der einzelnen Notationen der UB Hildesheim, so dass sie nicht in das erstellte alphabetische Register informationswissenschaftlicher Begriffe aufgenommen werden konnten und somit keine Zuweisung auf die Rubrik der Klassifikation möglich war. Teilweise brachte aber auch der Abgleich keine Treffer für Notationen, deren Titel in die Klassifikation aufgenommen werden sollten.

Da man vermuten durfte, dass sich für einige dieser Themengebiete dennoch Bücher in der UB Hildesheim finden lassen, wurden die Bezeichnungen dieser Themengebiete als Suchanfragen an den OPC4 geschickt. Aber auch andere Rubriken, bei denen schon Bücher im Bestand gefunden worden waren, wurden aus Gründen der Vollständigkeit durch Anfragen an den OPC4 ergänzt. Dabei wurde mit der Titelstichwörter- und Schlagwortsuche gearbeitet. Die Ergebnisse wurden in Dateien mit den Namen der Themengebiete abgespeichert und der Klassifikation zugewiesen (auftretende Dubletten wurden getilgt).

## **5 Konzept einer Zugriffsmöglichkeit auf informationswissenschaftliche Bücher über eine virtuelle Ordnungssystematik**

Mehrfach wurde im Verlauf dieser Arbeit darauf hingewiesen, wie sinnvoll es wäre, wenn Studierende der Informationswissenschaft in Hildesheim einen virtuellen Zugriff auf Literatur ihres Studienfaches hätten. Die Nachteile des bestehenden Systems wurden erläutert (siehe Kapitel 1.2.3 und Kapitel 2). Die für den Benutzer aus einer virtuellen Ordnungssystematik resultierenden Vorteile wären:

- Der Bestand an informationswissenschaftlicher Literatur stünde in gesamelter und übersichtlicher Form zur Verfügung. Studierende könnten sich durch Browsing einen schnellen Überblick über die geführten Bücher verschaffen.
- Einzelne Titel könnten unter mehreren Themenbereichen geführt werden. Auch ausgeliehene Titel und solche aus dem Magazin und den Semesterapparaten wären so präsent.
- Anpassungen auf Entwicklungen innerhalb der wissenschaftlichen Disziplin könnten schneller und problemloser erfolgen, da nur Veränderungen der virtuellen Inhalte erfolgen müssten.
- Bestände anderer Bibliotheken könnten dem Benutzer durch eine direkte Verlinkung in entsprechende Themenbereiche von informationswissenschaftlichen Klassifikationen anderer Standorte erschlossen werden.

Um den Benutzern der Hildesheimer Universitätsbibliothek einen Browsing-Zugang zu den ermittelten informationswissenschaftlichen Büchern zu ermöglichen, musste eine neue Klassifikation für diesen Literaturbestand erstellt werden und eine Zuordnung der Titel auf diese erfolgen.

### **5.1 Erarbeitung einer informationswissenschaftlichen Klassifikation**

Eine virtuelle Aufstellungssystematik für die informationswissenschaftliche Literatur in der UB Hildesheim sollte nach Möglichkeit die Vorteile und Stärken bereits bestehender Systematiken übernehmen und, wenn möglich, deren Nachteile und Mängel durch Verbesserungen vermeiden. Zu diesem Zweck wurde bereits eine Evaluierung der informationswissenschaftlichen Klassifikationen der Bibliotheken in Konstanz, Regensburg und Saarbrücken durchgeführt (siehe Kapitel 3.2).

Diese Klassifikationen dienten bei dem Entwurf einer neuen Systematik als Vorlage; die erwähnten Schaubilder (siehe Anlage 2, 3, 4) verhalfen dabei, eine schnelle Übersicht über die unterschiedlichen Gliederungspunkte zu gewinnen. Außerdem wurde die vom Hochschulverband für Informationswissenschaften aufgestellte Gliederung für informationswissenschaftliche Lehrveranstaltungen an verschiedenen Hochschulen herangezogen (siehe beiliegende CD-ROM). Eingang in die neue Systematik fanden ebenfalls Inhalte der Fachsystematik der Universitäts- und Landesbibliothek Düsseldorf (Bereich „Buch-, Bibliotheks- und Informationswesen“, Rubrik „Informationswissenschaft und Dokumentation“) und der „WebSite Methodik“ von der FH Darmstadt. Auch die im Verlauf der Arbeit entstandene Liste mit informationswissenschaftlichen Begriffen stand als Ideengeber und Orientierungshilfe zur Verfügung.

Den wohl größten Einfluss auf die Gestaltung der neuen Systematik hatten die Klassifikationen des kid und des HI. Die für die Konstanzer Systematik sprechenden Argumente wurden bereits genauer ausgeführt: Das kid stellt aufgrund der großen Menge an zu verwaltenden Titeln eine entsprechend ausgereifte Ordnungssystematik zur Verfügung. Der HI hat seine Klassifikation nach intensiven Beratungen bei einer Mitglieder-Tagung der Informationswissenschaftler erarbeitet.

In einem ersten Entwurf (siehe Anlage 15) wurde aus Gründen der Übersichtlichkeit versucht, die Anzahl der Gliederungspunkte in der obersten Ebene mög-

lichst gering zu halten. Dabei ergab sich eine Einteilung in die sechs Kategorien:

- Allgemeines
- Grundlagen / Theorie
- Informationstechnik / Elektronische Datenverarbeitung
- Informationsaufbereitung, -suche und -vermittlung
- Aspekte der Informationspraxis
- Anwendungsbereiche

Bei genauerer, gemeinsamer Betrachtung mit den beiden Betreuern stellte sich heraus, dass die drei ersten Hauptkategorien in dieser Form sinnvoll sind und übernommen werden können. Die drei letzten Punkte bedurften einer weiteren Überarbeitung. Durch die starke Konzentration würde für einen Benutzer so nicht auf den ersten Blick ersichtlich, welche Themengebiete sich hinter den Gliederungspunkten verbergen. Außerdem waren unter „Anwendungsbereiche“ Gebiete aufgelistet, bei denen es sich eigentlich um Kernbereiche der Informationswissenschaft handelt („Maschinelle Sprachverarbeitung“, „Computerlinguistik“). Das Fehlen von Begriffen wurde festgestellt (z.B. Informationsethik). Deshalb wurde entschieden, die vom HI erarbeitete und bis zu diesem Zeitpunkt noch nicht berücksichtigte Gliederung der Lehrveranstaltungen in den Aufbau der Klassifikation mit einzubeziehen und diese entsprechend umzustrukturieren.

Dementsprechend zahlreicher sind die Themenbereiche, die sich im zweiten Entwurf (siehe Anlage 16) in der ersten Gliederungsebene befanden. Aus den ehemals 6 waren jetzt 18 geworden:

- Allgemeines
- Grundlagen / Theorie
- Informationstechnik / Elektronische Datenverarbeitung
- Information und Gesellschaft / Informationspolitik
- Wissensvermittlung / Informations- und Dokumentationsstellen
- Informationsmanagement / Wissensmanagement
- Information Retrieval
- Informationssoziologie
- Informationsvisualisierung



- Informations- und Dokumentationssysteme
- Künstliche Intelligenz
- Informationsqualität und Evaluierung
- Sprachtechnologie
- Anwendungsbereiche der Informationswissenschaft
- Informationswirtschaft / Informationsmarkt
- Hypermedia / Multimedia
- Computer Mediated Communication
- Human Computer Interaction

Jedem der Gliederungspunkte sind Unterrubriken zugeordnet, die teilweise selbst weitere Unterrubriken besitzen. Insgesamt verfügt die gesamte Systematik über vier Ebenen (Endversion der Klassifikation: siehe beiliegende CD-ROM).

Der dritte Entwurf der Klassifikation (siehe Anlage 17) wies im Vergleich zu der vorhergehenden Version einige weitere Änderungen auf. So fiel die Oberkategorie „Informationssoziologie“ weg und die vorher darunter gefassten Unterrubriken wurden der Oberkategorie „Information und Gesellschaft / Informationspolitik“ zugewiesen. Auch die Kategorie „Hypermedia / Multimedia“ tauchte in dieser dritten Version nicht mehr als Oberkategorie auf, sondern wurde der Kategorie „Human Computer Interaction“ untergeordnet. Diese Veränderung wurde zugunsten einer genaueren Abgrenzbarkeit der Themengebiete vorgenommen, da der Begriff Multimedia für eine Kategorie der ersten Ebene eine sehr allgemeine Bezeichnung darstellt. Die zuvor unter „Hypermedia / Multimedia“ untergeordneten Sachgebiete wurden verteilt. „Authorware“ befand sich nun bei den Entwicklungsumgebungen unter dem ebenfalls neu gestalteten und dabei feiner gegliederten Thema „Programmiersprachen“, die vormals mit „Standards der Wissensdarstellung (UML, SGML, XML u.a.)“ bezeichnete Unterrubrik wurde als „Normen (UML, SGML, XML u.a.)“ unter „Grundlagen / Theorie“ klassifiziert. Eine Namensänderung erfuhr die Oberkategorie „Anwendungsbereiche der Informationswissenschaft“, die nun die Bezeichnung „Informationswissenschaftliche Anwendungen in anderen Sachgebieten“ trägt. Zusätzlich wurden die Kategorien in eine neue, nach Bedeutung gewichtete Reihenfolge gestellt (genau

wie auch die Rubriken der zweiten und dritten Ebene), die so u.a. den Schwerpunkten der Lehre an der Universität Hildesheim gerecht werden soll.

Im Vergleich zu diesen auffälligeren Umgestaltungen der neuen Systematik waren die weiteren Veränderungen von der zweiten zur dritten Version weniger bedeutsam. Es handelte sich dabei um Änderungen bei Rubriken der zweiten Ebene:

- das Verschieben von einer Rubrik in andere Oberkategorien („Expertensysteme“ von „Informations- und Dokumentationssysteme“ zu „Künstliche Intelligenz“, „Elektronische Enzyklopädien“ zu „Wissensvermittlung / Informations- und Dokumentationsstellen“, „Spracherkennung“ von „Künstliche Intelligenz“ zu „Sprachtechnologie“)
- das Wegfallen von Unterrubriken („Cliparts / Icons“ und „Videofilme“), die Ergänzung von Rubriken („digitale Fotografie“ durch „digitaler Film“)
- die Zusammenfassung von Rubriken („Maschinelle Sprachverarbeitung“ und „Spracherkennung“, „Data Mining“ und „Knowledge Discovery“)
- das Einfügen neuer Rubriken in die zweite Ebene („Informationsgesellschaft“ unter „Information und Gesellschaft / Informationspolitik“, „Internationales Informationsmanagement“ unter „Informationsmanagement / Wissensmanagement“)

Die in der dritten und vierten Ebene der Klassifikation vorgenommenen Veränderungen sind geringfügig. Lediglich die Rubriken „Korrektur-, Wandel-, Sortier- und Mischprogramme“ und „Fehlererkennung“ wurden entfernt, bei den „Rechnern“ kamen die „Großrechner“ hinzu, die „Verlage“ wurden um „E-Books und E-Journals“ erweitert und die „Programmiersprachen“ erhielten, wie schon weiter oben beschrieben, in einer vierten Ebene eine genauere Untergliederung.

Die neuen Oberkategorien der Klassifikation waren somit (siehe Anlage 17):

- Allgemeines
- Grundlagen / Theorie
- Informationstechnik / Elektronische Datenverarbeitung
- Information Retrieval
- Informationsqualität und Evaluierung

- Informationsmanagement / Wissensmanagement
- Informationswirtschaft / Informationsmarkt
- Human Computer Interaction
- Sprachtechnologie
- Computer Mediated Communication
- Informationsvisualisierung
- Wissensvermittlung / Informations- und Dokumentationsstellen
- Informations- und Dokumentationssysteme
- Künstliche Intelligenz
- Information und Gesellschaft / Informationspolitik
- Informationswissenschaftliche Anwendungen in anderen Sachgebieten

Diese dritte Version der neuen Klassifikation informationswissenschaftlicher Literatur enthält, wie die zweite Version auch, wieder vier Ebenen. Da bekanntermaßen eine Vielzahl von Ebenen dem User die Orientierung erschweren, taucht die vierte Ebene nur an drei Stellen in der Systematik auf und kann auch dort weggelassen werden, wenn sie sich als unübersichtlich erweisen sollten.

Die vierte und letzte Version der Systematik (siehe Anlage 18) entstand nach einer Besprechung mit Frau Müller und Herrn Pretz am 17.04.2002 (siehe Gesprächsprotokoll Müller / Pretz, Anlage 21), in der die dritte Version vorgestellt wurde. Hierbei wurde von beiden Anwesenden darauf aufmerksam gemacht, dass sich unter der Rubrik „Allgemeines“, die in jeder Oberkategorie angelegt ist, man Bücher erwarten könnte, die mehrere Themenbereiche dieser Kategorie der ersten Ebene in allgemeinerer, oberflächlicherer Form behandeln. Tatsächlich sollten dort aber Werke Eingang finden, die in anderen Rubriken nicht untergebracht werden können. Aus diesem Grund war es sinnvoll, diese Rubrik in „Sonstiges“ umzubenennen. Außerdem wurde die Systemstelle „Firmen“ von der Oberkategorie „Informationstechnik / Elektronische Datenverarbeitung“ in die Oberkategorie „Allgemeines“ verschoben, da an dieser Stelle klassifizierte Bücher alle anderen Wissensgebiete der Informationswissenschaft betreffen können.

## **5.2 Einordnung des Hildesheimer Bestandes in die neue Klassifikation**

### **5.2.1 Zuordnung der Hildesheimer Notationen zu der Klassifikation**

Nachdem wie beschrieben der informationswissenschaftliche Literaturbestand in der Universitätsbibliothek Hildesheim ermittelt wurde, sollten nun die einzelnen Titel den in der Klassifikation enthaltenen Themenbereichen zugeordnet werden. Dafür wurde in einem ersten Schritt wieder die erstellte Liste informationswissenschaftlicher Begriffe (siehe Kapitel 4.1) mit den zugeordneten Notationen der UB Hildesheim herangezogen.

Jedem der Themengebiete in der neuen Systematik wurden die in der Liste ermittelten Notationen zugeordnet. Bei diesem Arbeitsschritt wurden die Beschreibungen der einzelnen Notationen noch einmal genau untersucht, um die Übereinstimmung mit den Inhalten der Klassifikationsgliederung sicher zu stellen. Teilweise wurden dabei auch kleinere Korrekturen an der Gliederung vorgenommen, wenn eventuelle Überschneidungen bzw. Redundanzen offensichtlich wurden. Das Ergebnis wurde in einer Tabelle abgelegt (Auszug: siehe Anlage 19, komplette Tabelle: siehe beiliegende CD-ROM).

### **5.2.2 Zuweisung der gefundenen Titel auf die Klassifikation**

Damit die Hildesheimer Titel den einzelnen Rubriken der neuen Systematik zugeordnet werden konnten, erfolgte als Ergebnis des Parsens neben der Ausgabe des Autors und des Titels auch die Angabe der zugehörigen Signatur und der eingetragenen lokalen Sachgruppen (Nebenstellen). Die Titel mit Nebenstellen konnten somit mehreren Notationen zugeordnet werden und dadurch unter Umständen auch an mehreren Stellen in der Klassifikation Eingang finden. Auf diese Weise konnte erneut der Vorteil der Virtualität ausgenutzt werden. Die Titel einer gemeinsamen Notation mussten dafür aber zuerst in Dateien zusammen gefasst werden.

Für diesen Zweck wurde wiederum ein eigener Parser geschrieben, genau genommen für jede Notation ein eigener Parser (als Beispiel für die Notation CSC siehe `not_csc.awk`, Anlage 14). Darin wurden die Ergebnisdatei des Abgleichs und der Bücherbestellungen zeilenweise ausgelesen und es wurde jeweils ü-

berprüft, ob sich die gesuchte Notation in ihr befand, sowohl als Signatur als auch als Nebenstelle. Über eine For-Schleife wurden alle theoretisch möglichen Notationen eines Sachgebietes durchlaufen und die Treffer wurden jeweils in eine Datei mit dem Namen der enthaltenen Notation ausgegeben.

Nachdem nun Gruppen von Titeln mit gleichlautenden Notationen vorlagen, mussten diese auf die einzelnen Rubriken verteilt werden. Dafür diente als Basis die Tabelle, deren Erstellung in Kapitel 5.2.1 beschrieben wurde. Die Zuweisung erfolgte per Hand im Rahmen eines HTML-Dokumentes (siehe Kapitel 5.3). Die einzelnen Dateien mit den Notationen wurden dabei entsprechend ihrer Zugehörigkeit zu einzelnen Themengebieten der Klassifikation zusammengefasst und in einer neuen Datei unter dem Namen des Themengebieten abgespeichert. Diese Datei wurde dann mit der entsprechenden Rubrik in dem HTML-Dokument verlinkt.

Bei der Zuweisung zeigte sich, dass die vorher einer Rubrik zugeordneten Notationen nicht immer hundertprozentig den Themen der Sachgebiete entsprachen (siehe auch Kapitel 7). Oftmals mussten die Dateien der Notationen nach relevanten Titeln durchsucht werden, wenn sich auf den ersten Blick eine zu große thematische Heterogenität herausstellte. Einige zugewiesene Systemstellen wurden dabei aus der Klassifikations-Tabelle entfernt. Schon vorab waren einige Notationen gekennzeichnet worden, bei denen von vornherein keine völlige inhaltliche Kongruenz mit einer Klassifikationsrubrik bestand. Es handelte sich also bei dem beschriebenen Arbeitsschritt auch um eine stichpunktartige intellektuelle Prüfung der zugeordneten Titel. Allerdings muss berücksichtigt werden, dass aufgrund der enormen Menge von ca. 5.000 Büchern diese Kontrolle nur in beschränktem Umfang möglich war.

### **5.3 Zugriff über ein HTML-Dokument**

Aus Gründen der Anschaulichkeit und zur besseren Vermittlung des Ergebnisses der durchgeführten Arbeiten, wurde die neue Klassifikation in Form einer HTML-Struktur aufbearbeitet. Diese sollte in keinsten Weise gehobenen An-

sprüchen des Web-Designs entsprechen, Ziel war lediglich eine übersichtliche Darstellung. Die Grundstruktur wurde dabei mit dem Composer des Browsers Netscape Navigator 4.7 erstellt, ein Teil der notwendigen Arbeiten wurde mit Ultraedit, der auch als HTML-Editor geeignet ist, ausgeführt.

Die einzelnen Rubriken der neuen Klassifikation wurden untereinander verlinkt, so dass man durch die verschiedenen Dokumente navigieren kann (siehe beiliegende CD-ROM). Die obersten Kategorien befinden sich auf einer gemeinsamen Seite; jede einzelne Kategorie ist mit einem Link zu der Seite mit den zugehörigen Unterkategorien belegt (siehe Abbildung 14). Klickt man auf eine Kategorie der zweiten, dritten oder vierten Ebene, so kommt man direkt zu einer Datei, die die für dieses Themengebiet relevanten Titel zeigt. Von jeder Seite gelangt man wieder zurück auf die oberste Ebene.

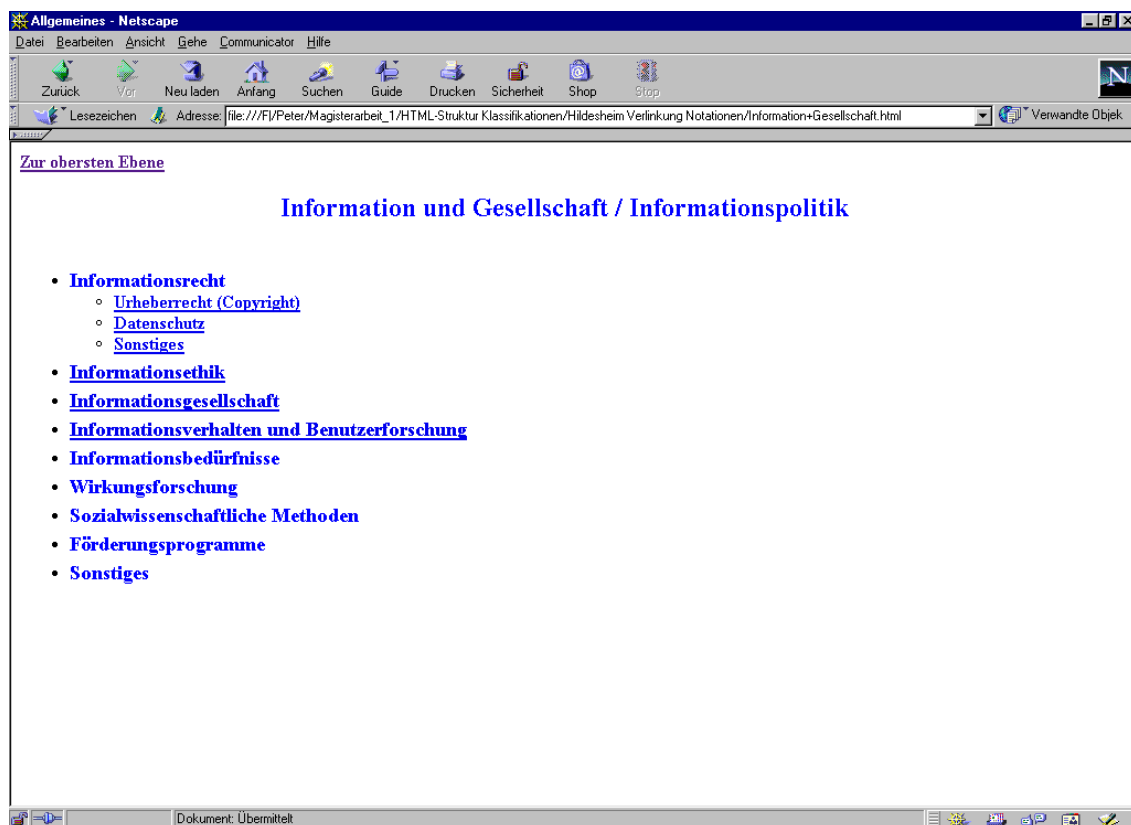


Abbildung 14: HTML-Seite „Information und Gesellschaft / Informationspolitik“

#### **5.4 Verlinkung einzelner Systemstellen der neuen Klassifikation mit Systemstellen anderer informationswissenschaftlicher Klassifikationen**

Schon zu Beginn der vorliegenden Masterarbeit war über die Möglichkeit nachgedacht worden, den Hildesheimer Benutzern der Universitätsbibliothek auch die Literatur an anderen Standorten verfügbar zu machen. Aus diesem Grund sollte anhand der neuen Systematik aufgezeigt werden, wo Querverbindungen zu Beständen anderer Bibliotheken aus demselben oder einem ähnlich gearteten Themenbereich bestehen. Es bot sich an, dies mit dem HTML-Dokument der neuen Klassifikation darzustellen.

Die Systematiken der Standorte Konstanz, Regensburg und Saarbrücken lagen ebenfalls als Webseiten vor. Somit konnte über eine Verlinkung von Themengebieten auf die entsprechende Position in der Klassifikation verwiesen werden. Dabei diene die Hildesheimer Klassifikation als Ausgangspunkt, d.h. ausgehend von einer Hildesheimer Rubrik wurde nach thematisch verwandten Rubriken an den anderen drei genannten Standorten informationswissenschaftlicher Studiengänge gefahndet. Durch Anklicken landet man direkt in der entsprechenden Klassifikation. Da auf jedes Themengebiet immer nur ein Link gesetzt werden kann, musste für jede der drei Klassifikationen eine eigene Verlinkung zur Hildesheimer Systematik erstellt werden (siehe beiliegende CD-ROM).

In Zukunft sollen diese Kongruenzen zwischen den Klassifikationen bei der Recherche nach Literatur ausgenutzt werden, um dem Benutzer eine größere Menge an Literatur zu offerieren, ohne dass sich dieser in die Systematik einer anderen Bibliothek einarbeiten muss. Bei der Suche in einer Rubrik der neuen Hildesheimer Klassifikation soll ihm dann beispielsweise als Ergänzung ein Verweis auf Bestände anderer Bibliotheken angezeigt werden. Durch einfaches Anklicken würde er dann direkt an der für ihn interessanten Stelle in der dortigen Systematik landen und könnte sich über das vorliegende Angebot informieren. Gleiches gilt natürlich auch für den Benutzer an einem Standort außerhalb Hildesheims: Auch er soll in seiner Klassifikation die Möglichkeit erhalten, ohne größere Mühe an den Bücherbestand einer anderen Bibliothek, also beispielsweise den der Hildesheimer UB, zu gelangen.

## 6 Evaluierung

Das im Verlauf der Arbeit erzielte Resultat, die Klassifikation in Form einer HTML-Struktur mit den dazugehörigen Titeln, wurde einer stichpunktartigen Prüfung unterzogen, um so eine Aussage über dessen Güte machen zu können. Hierbei ist allerdings keine abschließende Bewertung möglich, die erst eine Durchführung von Benutzertests zulassen würde. Es soll lediglich darum gehen, an Beispielen einen Einblick in die Qualität des Ergebnisses zu vermitteln.

Eine Möglichkeit der Evaluierung ist dadurch gegeben, dass man die Einordnung der klassifizierten Hildesheimer Bücher mit der der Konstanzer Datei vergleicht, aus welcher der Großteil der Werke stammt. Man kann so feststellen, unter welcher Rubrik ein Titel Eingang in die neue Systematik gefunden hat und wo er in der Konstanzer Klassifikation abgelegt ist. Auf diese Weise kann geprüft werden, wie gut das Resultat der halbautomatisch bzw. intellektuell vollzogenen Zuweisung der Titel im Vergleich zur Einordnung im Konstanzer kid ist, bei der jedes Buch von einem Bibliothekar einzeln auf seinen Bestimmungsort hin untersucht wird. Die Vermutung liegt nahe, dass Letzteres Erstem an Qualität überlegen sein dürfte.

Für die Untersuchung wurden 16 Titel verschiedener Themengebiete (aus allen Oberkategorien außer „Allgemeines“) per Zufallsprinzip aus der Hildesheimer Klassifikation ausgewählt. Dann wurde eine Suchanfrage an das Konstanzer Koala gestellt, um so den Standort im kid herauszufinden. Die dabei ermittelten Rubriken wurden gegenübergestellt. Diese Form der Evaluierung wurde durchgeführt, bevor die aus den Suchanfragen an den Hildesheimer OPAC stammenden Titel in die Klassifikation eingefügt worden waren, da bei diesen eine geringere Wahrscheinlichkeit bestand, dass sie auch im kid zu finden waren.

<b>Autor und Kurztitel</b>	<b>Hildesheimer Rubrik</b>	<b>Konstanzer Rubrik</b>
Behme: xml in der praxis: professionelles web-publishing mit der	Grundlagen / Theorie – Normen (UML, SGML, XML u.a.)	Elektronische Datenverarbeitung – HTML, CGI, XML, WML



<b>Autor und Kurztitel</b>	<b>Hildesheimer Rubrik</b>	<b>Konstanzer Rubrik</b>
Anahory: data warehouse: planung, implementierung und administration	Informationstechnik / EDV – Datenbanken – Sonstiges	Informationstechnologie – Software – analytische Informationssysteme, Datenbanksysteme, SIR, Data Warehouse
Dworatschek: personal-computer und datenschutz	Information und Gesellschaft / Informationspolitik – Informationsrecht – Datenschutz	Elektronische Datenverarbeitung – Datenschutz, Datensicherung, Computer-Viren
Baeza-yates: modern information retrieval	Information Retrieval – Sonstiges	Informationstechnologie – Software – Zugriffsverfahren, Information Retrieval, Suchverfahren, Downloading
Merx: qualitaetssicherung bei multimedia-projekten: mit 5 tabellen	Informationsqualität und Evaluierung – Sonstiges	Elektronische Datenverarbeitung – Datenbanken – Hypermedia, Multimedia, Authorware, Premo
Rau: integrierte buero-kommunikation: organisation und technik	Informationsmanagement / Wissensmanagement – Betriebliche Information & Kommunikation	Informations-Management, Wirtschaftsinformatik – Rationalisierung und Automatisierung in Büro – Büro und Verwaltung, office automation, Büro-kommunikation, Workflow-Management
Alpar: kommerzielle nutzung des internet: unterstuetzung von marketing,	Informationswirtschaft / Informationsmarkt – Elektronische Märkte	elektronische Datenverarbeitung – Internet, WWW, Eunet, Mosaic, Usenet, WAP
Fährnrich: software-ergonomie	Human Computer Interaction – Software-Ergonomie	Programmierung: Allgemeines und Theorie – Vorgang des Programmierens, Psychologie
Handke: sprachverarbeitung mit lisp und prolog auf dem pc	Sprachtechnologie – Maschinelle Sprachverarbeitung / Spracherkennung	Künstliche Intelligenz – Spracherkennung, speech recognition, Sprachverarbeitung
Berge: computer mediated communication and the online classroom	Computer Mediated Communication – Sonstiges Und: Wissensvermittlung / Informations- und Dokumentationsstellen – Computerbasiertes Lernen und Unterrichten	Sozialwissenschaften – Erziehungswissenschaft, Pädagogik, CAI = Computer assisted instruction, CAL = Computer assisted learning CBT = Computer based Training

<b>Autor und Kurztitel</b>	<b>Hildesheimer Rubrik</b>	<b>Konstanzer Rubrik</b>
Gonzalez: digital image processing	Informationsvisualisierung – Bildbearbeitung	Computergraphik, Grafik-Formate, GKS = graphisches Kernsystem
Küffner: computerlernen und autorensysteme	Wissensvermittlung / Informations- und Dokumentationsstellen – Computerbasiertes Lernen und Unterrichten	Sozialwissenschaften – Erziehungswissenschaft, Pädagogik, CAI = Computer assisted instruction, CAL = Computer assisted learning CBT = Computer based Training
Vetter: aufbau betrieblicher informationssysteme mittels	Informations- und Dokumentationssysteme – Informationssysteme verschiedener Sachgebiete - Sozialwissenschaften	Informations-Management, Wirtschaftsinformatik
Brause: neuronale netze: eine einfuehrung in die neuroinformatik	Künstliche Intelligenz – Neuronale Netze	neurale Computer, neuronale Netze
Döring: sozialpsychologie des internet: die bedeutung des internet	Information und Gesellschaft / Informationspolitik – Informationsgesellschaft	elektronische Datenverarbeitung – Internet, WWW, EUNET, Mosaic, Usenet, WAP
Reggia: computer-assisted medical decision making	Informationswissenschaftliche Anwendungen in anderen Sachgebieten – Mathematik und Naturwissenschaften	Mathematik und Naturwissenschaften – Medizin, Gesundheitswesen, BMDP

Tabelle 7: Vergleich der Einordnung von 16 Titeln in der Konstanzer und Hildesheimer Klassifikation

Die Auswertung der Tabelle zeigt, dass in 11 von 16 Fällen (Autoren: Behme, Anahory, Dworatschek, Baeza-yates, Rau, Handke, Berge, Gonzalez, Küffner, Brause, Reggia) die Titel in sehr ähnlich oder gleich lautenden Rubriken der jeweiligen Klassifikationen zu finden sind. In den anderen fünf Fällen (Merx, Alpar, Fähnrich, Vetter, Döring) ist die Aufstellung in den unterschiedlichen Themenrubriken der Konstanzer und Hildesheimer Klassifikation nachvollziehbar. Mindestens in zwei Fällen (Alpar, Döring) scheint eine Aufstellung in der Hildesheimer Rubrik allerdings sinnvoller, da dort die Thematik des Buches spezifischer widerspiegelt wird. Das Zustandekommen der unterschiedlichen Einordnung liegt selbstverständlich auch daran, dass die Benennung und Anzahl der einzelnen Kategorien und deren Unterrubriken in den beiden Klassifikationen nicht identisch ist. Somit kann man bei diesem Ergebnis wohl von ei-

nem relativ hohen Grad der Entsprechung bei der Zuordnung der Titel sprechen. Es rechtfertigt auch, zumindest für den Fall der Hildesheimer und der Konstanzer Systematik, die Verlinkung von gleich und ähnlich lautenden Rubriken der beiden Ordnungssystematiken (siehe Kapitel 5.4).

Eine weitere Form der Evaluierung wurde durchgeführt, indem mit den Namen von Hildesheimer Systemstellen Suchanfragen an den OPAC gestellt wurden, und anhand von zehn zufällig ausgewählten Titeln nachgeprüft wurde, ob sie der neuen Klassifikation hinzugefügt wurde (Die Durchführung fand ebenfalls vor dem Hinzufügen der bei den OPAC-Suchanfragen gefundenen Titel statt.). Auf diese Weise sollte, auch hier wiederum nur sehr stichpunktartig, ermittelt werden, wie gut die Erfassung der vorhandenen informationswissenschaftlichen Bücher in der UB Hildesheim gelungen ist.

<b>Titel</b>	<b>Hildesheimer Rubrik</b>
Hobert: Datenschutz und Datensicherheit im Internet : Interdependenz und Korrelation von rechtlichen Grundlagen und technischen Möglichkeiten	Nicht erfasst
Poetzsch: Information Retrieval : Einführung in Grundlagen und Methoden /	Information Retrieval – Modelle des Information Retrieval
Amtrup: Incremental speech translation	Sprachtechnologie – Maschinelle Sprachverarbeitung Und: Sprachtechnologie – Maschinelle und maschinengestützte Übersetzung
Kraus: Photoshop 5 : Perfekte Bildbearbeitung für MacOS und Windows	Informationsvisualisierung – Computergraphik, Grafik-Formate
Schuster: Digital Cash : Zahlungssysteme im Internet	Nicht erfasst
Carroll: Human-computer interaction in the new millennium	Human Computer Interaction – Sonstiges
Keil-Slawik: Digitale Medien und gesellschaftliche Entwicklung : Arbeit, Recht und Gemeinschaft in der Informationsgesellschaft	Nicht erfasst
Kmuche: Strategischer Erfolgsfaktor Wissen : content management: Der Weg zum erfolgreichen Informationsmanagement	Informationsmanagement / Wissensmanagement – Sonstiges
Scherer: Neuronale Netze : Grundlagen und Anwendungen	Künstliche Intelligenz – Lernende Systeme und Anpassungssysteme (incl.

	Neuronale Netze)
Steed: Web-based training	Nicht erfasst

Tabelle 8: Titel aus OPAC-Suchanfragen in der Hildesheimer Klassifikation

Das Ergebnis der Anfragen (sechs von zehn Titeln wurden gefunden) bestätigt die vor Durchführung des Abgleichs aufgrund der Literaturrecherche (siehe Kapitel 3.1) geäußerte Vermutung, dass nämlich bis zu ca. 70 Prozent des Hildesheimer informationswissenschaftlichen Bücherbestandes auf dem beschrifteten Weg erfasst werden können. Aufgrund der geringen Anzahl von Büchern ist die Tabelle nur bedingt aussagekräftig, vermittelt aber im Ansatz einen ungefähren Eindruck des erzielten Erfassungsgrades.

Interessant war auch herauszufinden, ob sich für den Benutzer im Vergleich zur Suchanfrage an den OPAC die Menge an präsentierter Literatur erhöht bzw. durch weitere Titel ergänzt wird. Dafür wurde beispielhaft für zwei Themenbereiche eine kombinierte Titel- und Schlagwort-Suchanfrage an den OPC4 gestellt. Ausgewählt wurden die beiden Sachgebiete der Klassifikation „Informationstheorie“ und „Software Engineering“.

Zuerst wurde der Suchbegriff in den OPC4 eingegeben. Das Ergebnis wurde mit dem bestehenden Parser bearbeitet und entsprechend formatiert (siehe Kapitel 4.3.1.3), damit ein Vergleich mit der durch den Abgleich ermittelten Datei möglich war. Der Texteditor Ultraedit bietet die Möglichkeit, Textdateien zeilenweise auf ihre Identität hin zu vergleichen. In einer Ausgabeoption kann man sich die Zeilen anzeigen lassen, in denen sich die beiden Dateien voneinander unterscheiden und in welcher Datei sich diese Zeilen befinden.

Für das Themengebiet „Informationstheorie“ ergab dieser Vergleich, dass die Anfrage über den OPC4 63 Titel (ohne Minimalsätze) als Ergebnis lieferte, die entsprechende Rubrik der Klassifikation enthielt 69 Titel. Die inhaltliche Gegenüberstellung zeigte, dass das Suchergebnis im OPC4 49 Titel lieferte, die sich nicht in der neuen Systematik wiederfinden. Auf der anderen Seite führt die Klassifikation 55 Titel, die über eine Suchanfrage nicht zu ermitteln waren.

Ähnlich ist das Resultat der Untersuchung für die Rubrik „Software Engineering“. Hier lieferte die Recherche über den Online-Katalog 310 Treffer, die Klassifikation enthält an dieser Stelle 271 Titel. Durch die OPC4-Suchanfrage erhält man im Vergleich zur Klassifikation 215 zusätzliche Titel, die Ordnungsys-tematik bietet 176 Titel, die man durch Titel- und Schlagwortsuche nicht findet.

Diese Zahlen machen eins deutlich: Suchanfragen nach Titel- und Schlagwörtern und die zusammengestellte Klassifikation stellen bei der Recherche eine sinnvolle Kombinationsmöglichkeit dar. Wie schon im Kapitel 2.1 dargelegt, werden über die Titel- und Schlagwortsuche nicht immer alle relevanten Titel gefunden. Auf der anderen Seite kann auch die Klassifikation keinen Anspruch auf Lückenlosigkeit erheben, da die Erfassung und Zuweisung auf halbautomatischem Weg erfolgte, was eine gewisse Vagheit beinhaltet.

Beim Überfliegen der durch die Klassifikation zusätzlich hinzukommenden Titel bestätigt sich dieser Eindruck. Nicht alle, aber doch der Großteil dieser Bücher haben einen engen oder etwas weiteren Bezug (unter „Informationstheorie“ finden sich einige Bücher zu „Kryptologie“) zum Thema der Rubrik. Der teilweise etwas weitere Bezug ergibt sich aus der Tatsache, dass die Qualität der Zuweisung in der neuen Klassifikation nicht die Güte einer bibliothekarischen Prüfung haben kann. Auf der anderen Seite kann auf diese Weise einem Benutzer aber ein möglichst umfassender Überblick über ein Themengebiet vermittelt werden.

Die Anzahl der für jede Systemstelle der Klassifikation ermittelten Bücher, sowohl über den Abgleich als auch über die OPAC-Suchanfragen, wurden in der Tabelle mit allen Ebenen der neuen Klassifikation festgehalten (Auszug: siehe Anlage 19, komplette Tabelle: siehe beiliegende CD-ROM), um so einen Eindruck von der Gesamtmenge der enthaltenen Titel zu erhalten. Die Zahl von insgesamt 9.524 zugewiesenen Titeln (incl. Mehrfachnennungen in unterschiedlichen Themengebieten) spricht für eine quantitativ gelungene Zuordnung der ermittelten Titel.

Zusätzlich zu den beschriebenen Untersuchungen wäre es sinnvoll, die Klassifikation einer noch eingehenderen Evaluierung zu unterziehen, um damit den Mehrwert für einen Benutzer zu verdeutlichen. Nur so wäre es wohl auch möglich, die Universitätsbibliothek Hildesheim von einer Übernahme der Ordnungssystematik in das bestehende System zu überzeugen. In dem Gespräch vom 17.04.2002 betonte Herr Pretz (siehe Anlage 21), dass sich der zusätzliche Arbeitsaufwand für die Bibliothekare nur nach einer genaueren Überprüfung des erzielten Ergebnisses rechtfertigen ließe. Dabei wäre u.a. zu untersuchen, wieviele der Bücher in der neuen Klassifikation an einer falschen Stelle eingeordnet sind und welche der Titel aus welchem Grund durch eine Suchanfrage nicht gefunden werden.

Herr Pretz wies in dieser Besprechung auch darauf hin, dass normalerweise bei der Erstellung einer Klassifikation beachtet werden sollte, dass Rubriken der gleichen Ebene die gleiche Gewichtung (gleiche Bedeutung für das Sachgebiet, ähnliche Menge an Titeln) haben. In dieser Arbeit wurde allerdings eher nach anwendungsbezogenen Kriterien vorgegangen, in dem man sich bei der obersten Ebene an den einzelnen Lehrgebieten der Informationswissenschaft orientierte, aus denen sich dann die untergeordneten Kategorien ergaben (siehe Anlage 21).

In demselben Gespräch mit Frau Müller und Herrn Pretz wurde auch auf eine weitere Problematik hingewiesen: Der Benutzer muss sich unter Umständen mit zwei Klassifikationen vertraut machen, der bereits bestehenden und der der Informationswissenschaft. Dabei könnten auch die bestehenden Doppelstellen (gleich lautende Teilgebiete in den verschiedenen Klassifikationen) problematisch sein, da sie die Eindeutigkeit in Frage stellen und für den Benutzer verwirrend sein können.

Zusammenfassend kann man jedoch festhalten, dass die Zuweisung der Titel auf die passenden Rubriken ein für die Art des Verfahrens gutes Ergebnis geliefert hat. Der Erfassungsgrad entspricht dabei vermutlich in etwa den vorab getätigten Prognosen. Die Benutzung der Klassifikation bei der Recherche stellt für den Benutzer insofern einen deutlichen Mehrwert dar, als er dort eine Viel-

zahl informationswissenschaftlicher Titel zu dem gewünschten Themengebiet findet, deren er auf anderem Weg nicht habhaft geworden wäre. Zu bemängeln bliebe lediglich, dass einigen Rubriken keine Titel zugewiesen werden konnten; dies ergibt sich daraus, dass manche informationswissenschaftliche Begriffe in der Beschreibung der Hildesheimer Notationen nicht auftauchen und somit keine Zuordnung einer Notation auf die betreffende Rubrik ausgeführt werden konnte (siehe Kapitel 5.2.1) bzw. dass keine sinnvolle Titel- oder Schlagwortsuche mit einem Rubriknamen der Klassifikation möglich war. Auf der anderen Seite wird es auch nicht für jedes Teilgebiet Titel in der UB Hildesheim geben. Die Klassifikation wurde bewusst so angelegt, dass auch zukünftig Literatur zu Themengebieten aufgenommen werden kann, die zur Zeit keine Schwerpunkte der Hildesheimer Lehre darstellen und aus diesem Grund auch bisher bei Bücherbestellungen nicht berücksichtigt wurden.

## 7 Ausblick

Nach den theoretischen Vorüberlegungen und der Problemverdeutlichung, den Recherchen und Evaluierungen, dem Auffinden eines Großteils der relevanten Literatur und der Erarbeitung einer eigenen Klassifikation mit Zuweisung der informationswissenschaftlichen Titel, womit die für diese Arbeit anvisierten Tätigkeiten abgeschlossen sind, soll am Ende noch ein Ausblick auf die weiteren Aufgaben erfolgen, die sich aus dem vorliegenden Resultat ergeben.

Bei der neuen Klassifikation mitsamt den zugeordneten Titeln handelt es sich gewissermaßen um eine Momentaufnahme. Informationswissenschaftliche Bücher, die neu in die UB Hildesheim eingestellt werden, müssen auf irgendeine Weise Eingang in diese Systematik finden. Eine wichtige Aufgabe wäre, Lösungen für diese Problematik zu entwerfen. Mögliche Ansätze sollen hier kurz aufgezeigt und grob umrissen werden.

Zum Einen käme eine automatische Zuordnung in Betracht, die den Vorteil hätte, dass sie einen relativ geringen Arbeitsaufwand erfordern würde. Als Orientierung müssten dabei wieder die den einzelnen Rubriken der Klassifikation zugeordneten Hildesheimer Notationen dienen. Die Aufnahme könnte entweder direkt bei der Neueinstellung eines Buches unter einer Notation erfolgen, die sich in der Klassifikations-Tabelle mit den zugeordneten Notationen wiederfindet (siehe Anlage 19). Oder sie könnte automatisch beim Ausleihen eines Buches erfolgen. Bei dieser Variante müsste beachtet werden, dass jedes Buch wirklich nur einmal unter jeder Rubrik erfasst wird.

Diese Art der Zuordnung beinhaltet allerdings einige Nachteile, auf die kurz eingegangen werden soll. Schon die in Kapitel 5.2.2 beschriebene Zuordnung der beim Abgleich ermittelten Titel hatte gezeigt, dass zwischen den Hildesheimer Notationen und den Themengebieten der Rubriken nicht immer eine völlige Kongruenz besteht. In vielen Fällen musste intellektuell geprüft werden, ob die Titel mit einer bestimmten Notation wirklich in die dafür vorgesehene Rubrik der



Klassifikation passten. Anhand von zwei Beispielen soll die Problematik verdeutlicht werden.

Die Notation BWL 687 beinhaltet Bücher zu den folgenden Themenbereichen: Betriebsinformatik / Informationssysteme [ Betriebsinformatik / Expertensysteme, Decision support; Betriebsinformatik / DFÜ, Internet, EDI; Electronic Commerce ]. Aufgrund dieser breit gefächerten Themenpalette wurde diese Notation mehreren Rubriken der Klassifikation zugewiesen (siehe beiliegende CD-ROM). Die stichpunktartige Überprüfung ergab auch, dass die Zuordnung größtenteils korrekt war, also die gefundenen Titel auf mehrere Rubriken der Systematik verteilt werden konnten. Dafür mussten aber die jeweils relevanten Titel selektiert werden; eine automatische Zuordnung jedes Titels in jede der zugewiesenen Rubriken hätte zu fehlerhaften Zuweisungen geführt.

Ähnliches gilt für das zweite Beispiel. Unter der Notation CSC 633 finden sich Bücher mit dem Inhalt: Information Retrieval [ Clustering; Query formulation; Retrieval models; Search process; Selection process ]. In der Klassifikation wurde diese Notation unter der Oberkategorie „Information Retrieval“ den Rubriken „Allgemeines“ und „Modelle des Information Retrieval“ zugeordnet. Auch hier war wiederum eine Selektion notwendig, um die jeweils relevanten Bücher herauszufinden.

Neben dieser Schwierigkeit muss beachtet werden, dass keineswegs alle Rubriken der Klassifikation mit Notationen der Hildesheimer Universitätsbibliothek bedient werden konnten. Für viele Rubriken findet sich in den Themenbeschreibungen der Notationen keine Entsprechung, so dass eine Zuordnung nicht möglich ist. Diese Bereiche würden somit bei einer automatischen Zuweisung leer ausgehen. Das Problem liegt einfach darin, dass nicht alle informationswissenschaftlichen Fachbegriffe, von denen einige neueren Datums sind, Eingang in die Systematik der UB Hildesheim gefunden haben.

Als Alternative zu dieser Erfassung hinzukommender informationswissenschaftlicher Bücher bietet sich die Einstellung per Hand an. Dabei müssten die Bibliothekare jedes neue Buch, das sie in die Klassifikation einordnen, mit der pas-

senden Notation versorgen. Zu diesem Zweck müsste vorab die gesamte neue Systematik mit einer zu bestimmenden Notation (beispielsweise ifw) ausgezeichnet werden. Die informationswissenschaftliche Klassifikation würde somit zum Bestandteil der bestehenden lokalen Ordnungssystematik der UB Hildesheim. Dieses Verfahren würde für die Bibliothekare nur einen geringfügig höheren Arbeitsaufwand bedeuten, da sie neben den anderen schon bestehenden Notationen eine weitere bei der Vergabe von Nebenstellen berücksichtigen müssten. Das Problem hierbei wäre allerdings, dass man auf irgendeine Weise den Altbestand der informationswissenschaftlichen Bücher mit den entsprechenden informationswissenschaftlichen Notationen versorgen und in die Datenbank einpflegen müsste, was vermutlich nur per Hand geschehen könnte. In jedem Fall wäre es eine wichtige Aufgabe, die Bibliothekare mit der neuen Systematik und ihren Besonderheiten vertraut zu machen.

Denkbar wäre auch, nur die neu hinzukommenden informationswissenschaftlichen Bücher mit einer Notation auszuzeichnen, und den ermittelten Altbestand unverändert zu lassen. Dann müsste eine Möglichkeit geschaffen werden, bei der der Benutzer beim Browsen in der Klassifikation die neuen Bücher über eine automatische Datenbankabfrage sehen kann, während ihm der im Rahmen dieser Arbeit ermittelte Altbestand sozusagen statisch angezeigt wird.

Ein weitere Idee für einen Einbau der neuen Klassifikation und die Einstellung hinzukommender Bücher stammt von Herrn Pretz (siehe Anlage 21). Er schlug vor, die beim Abgleich ermittelten Titel über ihre Signaturen den einzelnen Rubriken zuzuordnen. Beim Anklicken eines Teilgebietes der Klassifikation würde dann automatisch eine Suchanfrage nach den entsprechenden Signaturen gestartet. Neue Titel würden automatisiert eingestellt, indem bei bestimmten Suchanfragen (die genauer zu bestimmen wären) die Signaturen der Treffer automatisch der Rubrik zugeordnet würden. Auf diese Weise würde man es vermeiden, Einträge in die Datenbank vornehmen zu müssen.

Ein anderer in diesem Zusammenhang geäußelter Vorschlag von Herrn Pretz lautete dahingehend, dass ein Programm die Datenbank durchlaufen könnte und die Titel anhand der bestehenden Indexierung über Titel und Schlagwort

auf die neue Klassifikation zuordnet. Dies wäre dann auch für andere Fachbereiche möglich. Die dabei allerdings vorliegende Problematik bei der Titel- und Schlagwortsuche wurde bereits ausführlich erläutert (siehe Kapitel 2.1).

Auch die Klassifikation könnte noch verbessert werden. So müsste eine Klassifikationsbeschreibung mit Synonymen erstellt werden, bei der die Bezeichnungen der einzelnen Themenbereiche erläutert werden. Genauso sollte dem User eine Suchmöglichkeit angeboten werden, mit der er innerhalb der Klassifikationsbezeichnungen nach einem gewünschten Themengebieten forschen kann, ohne dass er sich erst mit der gesamten Systematik vertraut machen muss.

Weiter müssten sich an diese Arbeit anschließende Tätigkeiten das Ziel verfolgen, dem Benutzer den Zugriff auf die neue Systematik und die dahinter stehenden Bücher ermöglichen. Zum Einen könnte man Benutzertests durchführen, um zu sehen, wie gut die Klassifikation für das Suchverhalten der Studierenden geeignet und wie befriedigend die Trefferquote auf die Suchanfragen ist. Hierbei könnte auch verglichen werden mit den Ergebnissen durch Suchmöglichkeiten, die bisher schon zur Verfügung stehen (Anfrage mit Titelstichwörtern, Schlagwörtern usw.). Dies könnte im Rahmen der angesprochenen, ausführlicheren Evaluierung geschehen.

Zum Anderen müsste, um die Recherche mit der neuen Klassifikation im Rahmen des OPAC verfügbar zu machen, eine geeignete Benutzeroberfläche geschaffen werden, die die aktuellsten Erkenntnisse des Web-Designs berücksichtigt. Es wäre beispielsweise denkbar, dass dem Benutzer beim Betreten des OPAC durch einen Link eine spezifische Suche nach informationswissenschaftlicher Literatur angeboten wird. Er würde dann auf eine eigene Seite gelangen, wo ihm sowohl eine Eingabemöglichkeit für Suchanfragen, als auch die neue Systematik mit ihren Büchern zur Verfügung gestellt wird. Damit dieser Mehrwert für die Benutzer auch erkannt wird, wäre es allerdings wichtig, die Studierenden intensiv über das neue Angebot zu informieren.

## Literaturverzeichnis

**ACM:** Top Two Levels of The ACM Computing Classification System (1998).

<http://www.acm.org/class/1998/overview.html>

Verifizierungsdatum 19.09.2001

**Arnold, Guido** (2002): GNU ist nicht Unix!

<http://www.gnu.org/home.de.html>

Verifizierungsdatum 14.2.2002

**British Columbia Institute of Technology Library** : Library of Congress Classification System.

<http://www.lib.bcit.ca/LICO.htm>

Verifizierungsdatum 26.04.2002

**Chan, Lois Mai / Hodges, Theodora L.** (2000): The Library of Congress Classification. In: Marcella, Rita / Maltby, Arthur (Hrsg.): The Future of Classification. Aldershot: Gower, S. 105-119.

**FH Darmstadt<sup>1</sup>**, Fachbereich Informations- und Wissensmanagement: „WebSite Methodik“.

<http://www.iud.fh-darmstadt.de/methodik/#Start>

Verifizierungsdatum 26.04.2002

**FH Darmstadt<sup>2</sup>**, Fachbereich Informations- und Wissensmanagement: Thematische Karte von "WebSite 'Methodik'".

<http://www.iud.fh-darmstadt.de/methodik/search/themap.asp>

Verifizierungsdatum 26.04.2002

**Gemeinsamer Bibliotheksverbund:** Sachgebiete Basisklassifikation.

<http://gso.gbv.de:80/BCL>

Verifizierungsdatum 26.04.2002

**Gesellschaft für Informatik:** Fachbereiche der Gesellschaft für Informatik.

<http://www.gi-ev.de/wissenschaft/fachbereiche/index.html>

Verifizierungsdatum 30.01.2002

**GNU Project:** grep.

<http://www.gnu.org/software/grep/grep.html>

Verifizierungsdatum 26.04.2002

**Haller, Klaus** (<sup>3</sup>1998): Katalogkunde. Eine Einführung in die Formal- und Sacherschließung. München: K. G. Saur.

**McIlwaine, I. C.** (2000): UDC in the twenty-first century. In: Marcella, Rita / Maltby, Arthur (Hrsg.): The Future of Classification. Aldershot: Gower, S. 93-104.

**Mitchell, Joan S.** (2000): The Dewey Decimal Classification in the twenty-first century. In: Marcella, Rita / Maltby, Arthur (Hrsg.): The Future of Classification. Aldershot: Gower, S. 81-92.

**Oberhauser, Otto / Seidler, Wolfram** (2000): Reklassifizierung größerer fachspezifischer Bibliotheksbestände: Durchführbarkeitsstudie für die Fachbibliothek für Germanistik an der Universität Wien. Download von:

<http://www.bibliothek.uni-augsburg.de/allq/swk/sacher5.html>

Verifizierungsdatum 01.11.2001

**OCLC Forest Press**<sup>1</sup>: Introduction to the Dewey Decimal Classification.

[http://www.oclc.org/dewey/about/about\\_the\\_ddc.htm#moreinfo](http://www.oclc.org/dewey/about/about_the_ddc.htm#moreinfo)

Verifizierungsdatum 22.03.2002

**OCLC Forest Press**<sup>2</sup>: Summaries.

[http://www.oclc.org/dewey/about/ddc\\_21\\_summaries.htm](http://www.oclc.org/dewey/about/ddc_21_summaries.htm)

Verifizierungsdatum 22.03.2002

**Pica:** Über Pica.

<http://www.pica.nl/du/about/>

Verifizierungsdatum 20.05.2002

**Robbins, Arnold D.** (2001): GAWK: Effective AWK Programming. A User's Guide for GNU Awk. Edition 3. Download (gawk.html) von:

<http://www.gnu.org/manual/gawk/index.html>

Verifizierungsdatum 14.02.2002

**Rowley, Jennifer / Farrow, John** (<sup>3</sup>2000): Organizing Knowledge. An Introduction to Managing Access to Information. Aldershot: Gower.

**Stallman, Richard** (2001): The GNU Project.

<http://www.gnu.org/gnu/the-gnu-project.html>

Verifizierungsdatum 14.02.2002

**UDC Consortium:** About Universal Decimal Classification and the UDC Consortium.

<http://www.udcc.org/about.htm>

Verifizierungsdatum 23.03.2002

**Universitätsbibliothek Hildesheim**<sup>1</sup>: Web-OPAC.

<http://www.uni-hildesheim.de/UB/index.html>

Verifizierungsdatum 26.04.2002

**Universitätsbibliothek Hildesheim**<sup>2</sup>: Aufstellungssystematiken für die Freihandbestände.

[http://www.uni-hildesheim.de/UB/sysliste\\_2.html](http://www.uni-hildesheim.de/UB/sysliste_2.html)

Verifizierungsdatum 05.05.2002

**Universitätsbibliothek Hildesheim**<sup>3</sup>: Öffentlicher Online Katalog (OPAC), OPC4 Testinstallation, Verbundzentrale des GBV (VZG).

<http://hidbs2.bib.uni-hildesheim.de:8080/>

Verifizierungsdatum 27.04.2002

**Universitätsbibliothek Konstanz:** Systematische Gliederung „kid“. Kybernetik, Informatik, Datenverarbeitung und Informationswissenschaft.

<http://www.uni-konstanz.de/ZE/Bib/zs/kid-syst.htm>

Verifikationsdatum 26.04.2002

**Universitätsbibliothek Regensburg<sup>1</sup>:** Regensburger Verbundklassifikation – Online.

[http://www.bibliothek.uni-regensburg.de/rvko\\_neu/mytree.phtml](http://www.bibliothek.uni-regensburg.de/rvko_neu/mytree.phtml)

Verifizierungsdatum 24.03.2002

**Universitätsbibliothek Regensburg<sup>2</sup>:** Regensburger OPAC.

<http://rzbs5.bibliothek.uni-regensburg.de/webOPAC/welcomeDe.html>

Verifizierungsdatum 26.04.2002

**Universität Saarbrücken:** Klassifikationsschema der Institutsbibliothek Informationswissenschaft.

<http://www.phil.uni-sb.de/fr/infowiss/ib/klassif.html>

Verifikationsdatum 26.04.2002

**Universitäts- und Landesbibliothek Düsseldorf<sup>1</sup>:** Fachsystematik. Buch-, Bibliotheks- und Informationswesen.

<http://sun.ub.uni-duesseldorf.de/allhtml/system/bib2.htm>

Verifizierungsdatum 29.01.2002

**Universitäts- und Landesbibliothek Düsseldorf<sup>2</sup>:** Alphabetisches Register zu den Beständen des Studiengangs Informationswissenschaft.

<http://www.uni-duesseldorf.de/WWW/ulb/iudreg.html>

Verifizierungsdatum 10.10.2001

## Anhang

Anlage 1	Auszug aus dem Ergebnis der Literaturrecherche
Anlage 2	Regensburger Verbundklassifikation (RVK)
Anlage 3	Saarbrücker Klassifikation
Anlage 4	Konstanzer Klassifikation kid
Anlage 5	Auszug aus der Liste informationswissenschaftlicher Begriffe mit zugeordneten relevanten Hildesheimer Notationen
Anlage 6	Parser pars_k1.awk
Anlage 7	Parser pars_k2.awk
Anlage 8	Parser pars_k3.awk
Anlage 9	Parser pars_k4.awk
Anlage 10	Parser pars_k5.awk
Anlage 11	Parser pars_k6.awk
Anlage 12	Parser pars_h08.awk
Anlage 13	Parser pars_kh9.awk
Anlage 14	Parser not_csc.awk
Anlage 15	Hildesheimer Klassifikation (Entwurf 1 vom 15.01.2002)
Anlage 16	Hildesheimer Klassifikation (Entwurf 2 vom 20.02.2002)
Anlage 17	Hildesheimer Klassifikation (Entwurf 3 vom 08.03.2002)
Anlage 18	Hildesheimer Klassifikation (Entwurf 4 vom 22.04.2002)
Anlage 19	Auszug aus der Hildesheimer informationswissenschaftlichen Klassifikation mit den zugeordneten Notationen
Anlage 20	Protokoll des Gespräches mit Edwin Pretz vom 04.12.2001
Anlage 21	Protokoll des Abschlußgespräches zur Magisterarbeit mit Ansprechpartnern der Bibliothek (vom 17.04.2002)



Hiermit erkläre ich, die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel verfasst zu haben.

Hildesheim, den 30.05.2002

Peter Hanke

## Anhang

Anlage 1	<i>7.1.1 Auszug aus dem Ergebnis der Literaturrecherche</i>
Anlage 2	Regensburger Verbundklassifikation (RVK)
Anlage 3	Saarbrücker Klassifikation
Anlage 4	Konstanzer Klassifikation kid
Anlage 5	Auszug aus der Liste informationswissenschaftlicher Begriffe mit zugeordneten relevanten Hildesheimer Notationen
Anlage 6	Parser pars_k1.awk
Anlage 7	Parser pars_k2.awk
Anlage 8	Parser pars_k3.awk
Anlage 9	Parser pars_k4.awk
Anlage 10	Parser pars_k5.awk
Anlage 11	Parser pars_k6.awk
Anlage 12	Parser pars_h08.awk
Anlage 13	Parser pars_kh9.awk
Anlage 14	Parser not_csc.awk
Anlage 15	Hildesheimer Klassifikation (Entwurf 1 vom 15.01.2002)
Anlage 16	Hildesheimer Klassifikation (Entwurf 2 vom 20.02.2002)
Anlage 17	Hildesheimer Klassifikation (Entwurf 3 vom 08.03.2002)
Anlage 18	Hildesheimer Klassifikation (Entwurf 4 vom 22.04.2002)
Anlage 19	Auszug aus der Hildesheimer informationswissenschaftlichen Klassifikation mit den zugeordneten Notationen
Anlage 20	Protokoll des Gespräches mit Edwin Pretz vom 04.12.2001
Anlage 21	Protokoll des Abschlußgespräches zur Magisterarbeit mit Ansprechpartnern der Bibliothek (vom 17.04.2002)

7.1.1.1.1 Themengebiete	7.1.1.2 Signaturen und Notationen			
	Hildesheim	Konstanz	Regensburg	Saarbrücken
<b>E-Commerce (Hildesheim)</b>				
Electronic business : Methoden, Werkzeuge, Techniken und Systeme für den Unternehmenserfolg im Internet, 2000	BWL 687 : D17		QR 760	
Internet economics, 1998	CSC 645 : I54	wrr 495.f/ t97	440/QR 760 M159	
Principles of internet marketing, 2000	BWL 589 : H05		<ul style="list-style-type: none"> <li>• 406/QR 760 H251</li> <li>• QP 650</li> </ul>	
E-Commerce und Online-Marketing : Chancen, Risiken und Strategien, 2000	BWL 687 : K62		<ul style="list-style-type: none"> <li>• F 03/QP 650 K91(2)+1</li> <li>• QR 700</li> </ul>	200-4422
Handbook on electronic commerce : with 43 tables, 2000	BWL 687 : H05	wrp 238/ h16	<ul style="list-style-type: none"> <li>• F 03/QP 650 S353+1</li> <li>• 40/QR 760 S535</li> <li>• ST 370</li> </ul>	
Electronic commerce : Anwendungsbereiche und Potentiale der digitalen Geschäftsabwicklung, 2000	BWL 687 : E45	kid 836/ t46	<ul style="list-style-type: none"> <li>• F 03/QP 650 T465(2)+1</li> <li>• 40/QR 760 T465(2)</li> </ul>	250-4404
Der Internet-Guide fuer Manager, 1997	BWL 687 : R53		<ul style="list-style-type: none"> <li>• 406/QP 345 R738</li> <li>• ST 201</li> </ul>	
Information und Kommunikation : Technik und Anwendung in Wirtschaft und Medien, 1998	CSC 640 : B34			
Internet-Marketing und Electronic Commerce : Grundlagen - Rahmenbedingungen - Instrumente ; mit Erfolgsbeispielen, 2000	BWL 589 : F65	wrp 328/ f74a	<ul style="list-style-type: none"> <li>• 40/QP 650 F919 I6</li> <li>• QR 700</li> </ul>	
Kommerzielle Nutzung des Internet, 1998	CSC 645 ; BWL 687 ; CSC 221	<ul style="list-style-type: none"> <li>• kid 228.50:xh/a56</li> <li>• wrp 238/a56(2)</li> </ul>	<ul style="list-style-type: none"> <li>• F 03/WIR 242 314(2)</li> <li>• QP 345</li> <li>• ST 200</li> </ul>	

Anlage 1: Auszug aus dem Ergebnis der Literaturrecherche

**Buch- und Bibliothekswesen, Informationswissenschaft**

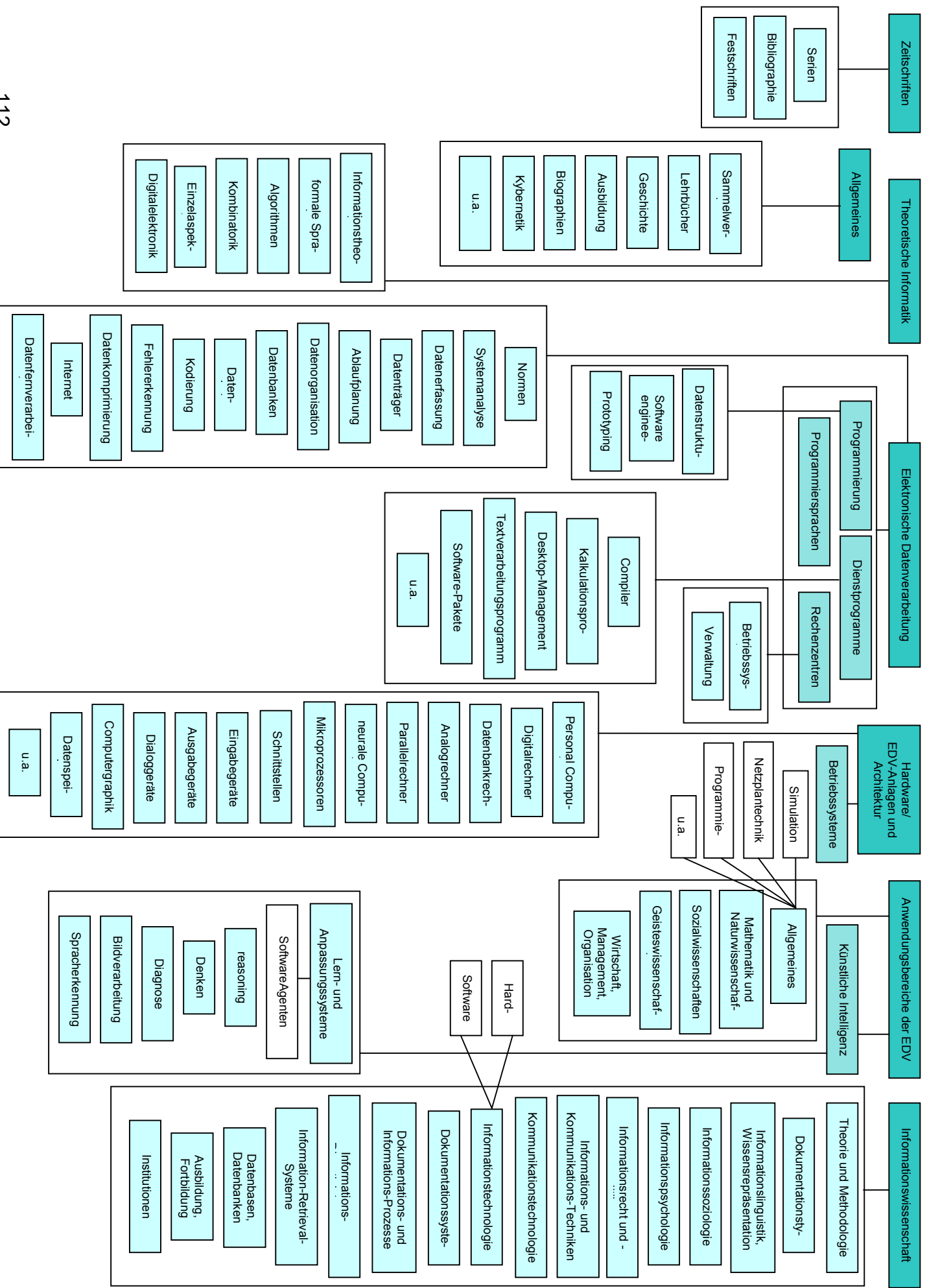
Informationswissenschaft

Bibliographien, Sammelchriften	Organisation, Beruf, Forschung	Nachschlagewerke, Allgemeine Darstellungen	Grundlagen, Theorie	Thesauri, Begriffsverzeichnisse	Informationspraxis	Informationsvermittlung, Informationsplatz	Fachliche Information und Dokumentation
Bibliographien periodisch	Institutionen, Verbände	Adressbücher	Informationstheorie, Kybernetik	Allgemeines, Wissenschaft, Hochschulwesen	Bibliographische Erfassung, Suchpraxis	Allgemeines	Allgemeines, Wissenschaft, Hochschulwesen
Bibliographien nichtperiodisch	Einrichtung und Ausstattung von Informationsstellen	Fachwörterbücher einsprachig	Semantik	Theologie, Religionswissenschaft	Referieren, Klassifizieren, Indexieren	Herstellung von Bibliographien und Informationsdiensten	Theologie, Religionswissenschaft
Fortschrittsberichte	Organisationsdiagnostik	Fachwörterbücher mehrsprachig	Ordnungslehre, Systematik	Philosophie, Psychologie	Selektive Informationsvermittlung (SDI)	Auskunft und Beratung	Philosophie, Psychologie
Schriftenreihen	Informationsnetze, Informationsverbund	Normen	Terminologie	Pädagogik	Automatisches	Wissensverbreitung, Informationsfluß	Pädagogik
Tagungs- und Kongreßberichte periodisch	Ausbildung	Lexika	Klassifikation	Sprach- und Literaturwissenschaft, Übersetzungswesen	Speicherung von Informationen	Nutzung	Sprach- und Literaturwissenschaft, Übersetzungswesen
Fest- und Gedendriften	Berufs- und Personalfragen	Handbücher	Klassifikationsvergleich	Volks- und Volkskunde	Technische Verfahren und Hilfsmittel	Informationsbedarf	Volks- und Volkskunde
Sammelwerke	Forschung	Einführungen, Abrisse	Deskription	Archäologie, Kunst	Kosten	Rehabilität	Archäologie, Kunst
		Gesamtdarstellungen	Thesauruskunde	Musik, Theater, Film	Bewertung, Effizienz, Qualität		Musik, Theater, Film
		Darstellungen zu mehreren Gebieten	Regeln für Thesauren	Politologie, Soziologie, Presse, Rundfunk	Mikroformen und Lesegeräte	Informations- und Dokumentationssysteme	Politologie, Soziologie, Presse, Rundfunk
		Geschichte der Informationswissenschaft	Klassifikationssysteme	u.a.	Sacherschließung	nach Art, geographisch	u.a.

## Anlage 2: Regensburger Verbundklassifikation (RVK)



## 8 Anlage 4: Konstanzer Klassifikation kid



**Anlage 5:** Auszug aus der Liste informationswissenschaftlicher Begriffe mit zugeordneten relevanten Hildesheimer Notationen

Software Engineering	CSC 320, CSC 321, CSC 322, CSC 323, CSC 325, CSC 326, CSC 327, CSC 328, CSC 329
Software-Ergonomie	CSC 309,
Software-Psychologie	CSC 309
Software-Management	CSC 943
speech recognition	Keine Treffer
Storage	CSC 132, CSC 333, CSC 352, CSC 731,
Sprachanalyse	Keine Treffer
Spracherkennung	LIN 460,
Sprachsynthese	LIN 460
Sprachverarbeitung	LIN 430
Suchmaschinen	Keine Treffer
Suchpraxis	Keine Treffer
Suchverfahren	CSC 633
Systemarchitektur	CSC 200, CSC 209,
Systematik	Keine Treffer
Systemverwaltung	CSC 944
Tabellenkalkulation	CSC 810
<i>Technische Dokumentation</i>	<i>LIN 263</i>
Telekommunikation	TFF 087
Teleconferencing	Keine Treffer
Teletext	Keine Treffer
Terminologie	LIN 280, LIN 281, LIN 285
Text processing	CSC 754
Textverarbeitungsprogramme	CSC 771
<i>Textlinguistik</i>	<i>GER 116, LIN 116,</i>
Thesauri	BUB 146, CSC 631
Transaktionssysteme	Keine Treffer
Übersetzungswissenschaft	LIN 310, LIN 315, LIN 319
UML = Unified Modeling Language	Keine Treffer
UNIX	CSC 363
Urheberrecht	REC 288, CSC 931,
USB	Keine Treffer
Usenet	Keine Treffer
user interface	CSC 322, CSC 652,

Notationen von kursiv gedruckten Begriffen wurden der neuen Klassifikation nicht zugewiesen; zwar fanden sie sich in der Klassifikationsbeschreibung der UB Hildesheim, eine genauere Prüfung der Inhalte ergab jedoch eine fehlende Relevanz.

```
BEGIN {RS = "#"}

# legt „#“ als neuen Separator zwischen den einzelnen Records fest,
# eröffnet somit an dieser Stelle eine neue Zeile

while (getline < "kid.txt" > 0)
{
    # Eingabefile lesen, solange noch Zeilen vorhanden sind
    zeile = $0
    # $0 steht für eine gesamte Zeile, ihr wird die Variable "zeile" zugewiesen
    print zeile >> "roh_ko.txt"
    # Ausgabe der Zeilen in die Datei roh_ko.txt
}

}
```



```
BEGIN {print "Konstanzer Autoren und Titel:"
main()
}
function main()
{
while (getline < "roh_ko.txt" > 0 )
{
if ( index($1, "kid") == 0 &&
index($2, "kid") == 0 &&
index($0, "19") > 0)
{
autorundtitel = substr($0, 1, (index($0, "19"))-1)
print autorundtitel >> "Tite_ko.txt"
}
# Ausschneiden des relevanten Strings
# Ausgabe in Datei Tite_ko.txt

if ( index($1, "kid") == 0 &&
index($2, "kid") == 0 &&
index($0, "20") > 0 &&
index($0, "19") == 0 &&
index($0, "9999") == 0)
{
autorundtitel = substr($0, 1, (index($0, "9999"))-1)
print autorundtitel >> "Tite_ko.txt"
}
# Ausschneiden des relevanten Strings
# Ausgabe in Datei Tite_ko.txt

# Auffinden des Autors und Titels bis zur Jahreszahl
# Auffinden des Autors und Titels bis zur Jahreszahl
}
```

```

    {
        autorundtitel = substr($0, 1, (index ($0, "20"))-1)
        print autorundtitel >> "Tite_ko.txt"

        # Ausschneiden des relevanten Strings
        # Ausgabe in Datei Tite_ko.txt
    }

    if ( index($1, "kid") == 0 &&
        index($2, "kid") == 0 &&
        index($0, "0") > 0 &&
        index($0, "19") == 0 &&
        index($0, "9999") == 0 &&
        index($0, "20") == 0 )
    {
        # Auffinden des Autors und Titels bis zur Jahreszahl

        autorundtitel = substr($0, 1, (index ($0, "0"))-1)
        print autorundtitel >> "Tite_ko.txt"

        # Ausschneiden des relevanten Strings
        # Ausgabe in Datei Tite_ko.txt
    }

    # Ende der while-Schleife
}

```

```

BEGIN {print "Konstanzer Endversion1:"
      main()
}

function main()
{
  while (getline < "Tite_ko.txt" > 0 )      # Eingabefile lesen, solange noch Zeilen vorhanden sind
  {
    autorundtitel = substr($0, 1, 55)        # Festlegen der Zeilenlänge

    if ( index(autorundtitel, "(") ==0)      # Titel ohne aufgehende Klammer suchen
    {
      print autorundtitel >> "Ti_ko.txt"
    }

    if ( index(autorundtitel, "(") >0 &&      # Titel mit aufgehender
        index(autorundtitel, ")") >0)        # und schließender Klammer suchen
    {
      print autorundtitel >> "Ti_ko.txt"
    }

    if ( index(autorundtitel, "(") >0 &&      # Titel mit aufgehender
        index(autorundtitel, ")") ==0)      # und ohne schließende Klammer suchen
    {
      #Wörter bis zur aufgehenden Klammer herauserschneiden
      titelohneklammer = substr(autorundtitel, 1, (index (autorundtitel, "(")-1)
      print titelohneklammer >> "Ti_ko.txt"
    }
  }
}
}

```

```

BEGIN {print "Konstanzer Autoren und Titel:"
      main()
}

function main()
{
  while (getline < "Kidnot.txt" > 0 )      # Eingabefile lesen, solange noch Zeilen vorhanden sind
  {
    titel = ""                             # Variable „titel“ auf Null setzen
    verfasser = ""                         # Variable „verfasser“ auf Null setzen
    es_gibt_einen_verfasser = 0           # Variable „es_gibt_einen_verfasser“ auf Null setzen

    while ((getline < "Kidnot.txt" > 0) && ( index ($0, "_" ) == 0 ) )  # Eingabefile bis Zeile mit Unterstrich auslesen
    {
      if ( index($0, "TIT") > 0 )
      {
        titel = substr ( $0, (index($0, "TIT")+8))      # Erfassen der relevanten Zeileninhalte
      }

      if ( index($0, "PER") > 0 && es_gibt_einen_verfasser == 0 )
      {
        verfasser = substr ( $0, (index($0, "PER")+8), (index($0, ",")-9))  # Erfassen der relevanten Zeileninhalte
        es_gibt_einen_verfasser = 1      # bewirkt, dass nur der jeweils erste Verfasser eines Titels erfasst wird
      }

      } # Ende der ersten while-Schleife

    } # Ende der zweiten while-Schleife
  }
}

```

```

BEGIN {print "Konstanzer Titel:"
      main()
}

function main()
{
    while (getline < "Titel_kono.txt" > 0 )    # Eingabefile lesen, solange noch Zeilen vorhanden sind
    {
        autorundtitel = substr($0, 1, 55)      # Festlegen der Zeilenlänge

        if ( index(autorundtitel, "(") == 0 )    # Titel ohne aufgehende Klammer suchen
        {
            print autorundtitel >> "Titel_kono.txt"
        }

        if ( index(autorundtitel, "(") > 0 &&      # Titel mit aufgehender
              index(autorundtitel, ")") > 0 )    # und schließender Klammer suchen
        {
            print autorundtitel >> "Titel_kono.txt"
        }

        if ( index(autorundtitel, "(") > 0 &&      # Titel mit aufgehender
              index(autorundtitel, ")") == 0 )    # und ohne schließende Klammer suchen
        {
            # Wörter bis zur aufgehenden Klammer herauserschneiden
            titelohneklammer = substr(autorundtitel, 1, (index (autorundtitel, "(")-1)
            print titelohneklammer >> "Titel_kono.txt"
        }
    }
}

```

```

BEGIN {print "Konstanzer Endversion2:"
      main()
}

function main()
{
    while (getline < "Ti_kon.txt" > 0 )      # Eingabefile lesen, solange noch Zeilen vorhanden sind
    {
        autorundtitel = $0

        if ( index (autorundtitel, "[") ==0) # Titel ohne eckige aufgehende Klammer suchen
        {
            print autorundtitel >> "Ti_kono.txt"
        }

        if ( index (autorundtitel, "[") >0 &&    # Titel mit aufgehender
            index (autorundtitel, "]") >0)      # und schließender eckiger Klammer suchen
        {
            print autorundtitel >> "Ti_kono.txt"
        }

        if ( index (autorundtitel, "[") >0 &&    # Titel mit aufgehender
            index (autorundtitel, "]") ==0)      # und ohne schließende eckige Klammer suchen
        {
            # Wörter bis zur aufgehenden eckigen Klammer herauserschneiden
            titelohneklammer = substr (autorundtitel, 1, (index (autorundtitel, "[")-1)
            print titelohneklammer >> "Ti_kono.txt"
        }
    }
}

```

```

BEGIN {print "Hildesheimer Autoren und Titel:"
      main()
}

function main()
{
    while (getline < "hilde.txt" > 0)      # Eingabefile lesen, solange noch Zeilen vorhanden sind
    {
        titel = ""
        verfasser = ""
        signatur = ""
        sachgruppen = ""
        es_gibt_einen_verfasser = 0      # alle Variablen auf Null setzen

        while ((getline < "hilde.txt" > 0) && ($1 != "Ausleihstatus:"))    # die Datei solange zeilenweise auslesen,
            # bis eine Zeile mit „Ausleihstatus:“ beginnt
        {
            if ( index($0, "Titel:") > 0 )      # Zeilen mit „Titel:“ suchen ($0 steht für die gesamte Zeile)
            {
                if ( index($0, "/" ) > 0 )      # Zeilen mit „/“ suchen
                {
                    titel = substr ( $0, (index($0, "Titel:")+7), (index($0, "/")-15))
                    # 15, da die maximale Anzahl von Leerschritten vor Titel vermutlich 8 beträgt,
                    # (8+7=15)
                }
            }
            else
            {
                titel = substr ( $0, (index($0, "Titel:")+7))
            }
        }
    }
}

```

```

if ( index($0, "Werk:") > 0 )           # Zeilen mit „Werk:“ suchen
{
    if ( index($0, "/" ) > 0 )
    {
        titel = substr ( $0, (index($0, "Werk:")+6), (index($0, "/")-19))
        # Bezeichnung „titel“ bleibt, da entweder "Titel" oder "Werk" erscheint und nicht beides
        # 19, da die maximale Anzahl von Leerschritten vor Werk 13 beträgt,
        # (13+6=19)
    }

    else
    {
        titel = substr ( $0, (index($0, "Werk:")+6))
    }
}

# Auffinden des Elementes „Verfasser:“ und Bestimmung des Nachnamens
if ( index($0, "Verfasser:") > 0 && index($0, ",") > 0 && index($0, "*") == 0 )
{
    if ($4 == ",")
    {
        verfasser = substr( $3, 1 ) # genauen Namen vor dem „“ ermitteln = erster Nachname
    }

    else
    {
        verfasser = substr( $4, 1 ) # genauen Namen vor dem „“ ermitteln = erster Nachname
    }

    es_gibt_einen_verfasser = 1      # bewirkt, dass nur der erste Verfasser erfasst wird
}

```



```

# durch das Element „*“ verändert sich die Position des Verfasser Namens in der Zeile
if ( index($0, "Verfasser:") > 0 && index($0, ",") > 0 && index($0, ".*") > 0)
{
    if ($5 == ",")
    {
        verfasser = substr( $3, 1) # genauen Namen vor dem „.“ ermitteln = erster Nachname
    }
    else
    {
        verfasser = substr( $4, 1) # genauen Namen vor dem „.“ ermitteln = erster Nachname
    }
    es_gibt_einen_verfasser = 1
}

# Auffinden des Elementes „Verfasser:“ und Bestimmung des Nachnamens
if ( index($0, "Verfasser:") > 0 && index($0, ",") == 0 && index($0, ".*") == 0)
{
    verfasser = substr( $NF, 1)      # $NF zeigt auf das letzte Feld (Wort) der Zeile
    es_gibt_einen_verfasser = 1
}

# Auffinden des Elementes „Verfasser:“ und Bestimmung des Nachnamens
if ( index($0, "Verfasser:") > 0 && index($0, ",") == 0 && index($0, ".*") > 0)
{
    verfasser = substr( $(NF-1), 1)
    es_gibt_einen_verfasser = 1
}

```

```
# Auffinden des Elements „Beteiligt“ und Bestimmung des Nachnamens
if ( index($0, "Beteiligt:") > 0 && index($0, ",") > 0 && es_gibt_einen_verfasser == 0 && index($0, "**") == 0 )
{
    if ($4 == ",")
    {
        verfasser = substr( $3, 1)      # genauen Namen vor dem „“ ermitteln = erster Nachname
    }
    else
    {
        verfasser = substr( $4, 1)      # genauen Namen vor dem „“ ermitteln = erster Nachname
    }
}

# durch das Element „“ verändert sich die Position des Verfasser Namens in der Zeile
if ( index($0, "Beteiligt:") > 0 && index($0, ",") > 0 && es_gibt_einen_verfasser == 0 && index($0, "**") > 0 )
{
    if ($5 == ",")
    {
        verfasser = substr( $3, 1)      # genauen Namen vor dem „“ ermitteln = erster Nachname
    }
    else
    {
        verfasser = substr( $4, 1)      # genauen Namen vor dem „“ ermitteln = erster Nachname
    }
}
```

```
# Auffinden des Elementes „Beteiligt“ und Bestimmung des Nachnamens
if ( index($0, "Beteiligt:") > 0 && index($0, ",") == 0 && es_gibt_einen_verfasser == 0 && index($0, "***") == 0 )
{
    verfasser = substr( $NF, 1)      # $NF zeigt auf das letzte Feld (Wort) der Zeile
}

# Auffinden des Elements „Beteiligt“ und Bestimmung des Nachnamens
if ( index($0, "Beteiligt:") > 0 && index($0, ",") == 0 && es_gibt_einen_verfasser == 0 && index($0, "***") > 0 )
{
    verfasser = substr( $(NF-1), 1)
}

if (index($0, "Sachgruppen:") > 0 && index($0, ".") == 0)
{
    sachgruppen = substr ($0, (index($0, "Sachgruppen:")+13))
}

if (index($0, "Signatur:") > 0)
{
    signatur = substr ($0, (index($0, "Signatur:")+10))
}

} # Ende der ersten while-Schleife

printf "%s: %s | %s | %s\n", verfasser, titel, signatur, sachgruppen >> "Ti_hi.txt"
# %s dient als Platzhalter für die Variablen "verfasser, titel, signatur, sachgruppen"

} # Ende der zweiten while-Schleife

}
```

```

BEGIN {print "Endergebnis:"
      main()
}

function main()
{
  while (getline <"Ti_koges.txt" > 0 )      # aus Konstanzer Gesamtfile lesen
  {
    konstanzer_titel = $0
    konstanzerautor = $1
    konstanzerwort1 = $2
    konstanzerwort2 = $3
    konstanzerwort3 = $4

    # die einzelnen Felder werden mit Variablen belegt

    while (getline <"Ti_hi.txt" > 0 )      # aus Hildesheimer File lesen
    {
      hildesheimertitel = $0
      hildesheimerautor = $1
      hildesheimerswort1 = $2
      hildesheimerswort2 = $3
      hildesheimerswort3 = $4

      # die einzelnen Felder werden mit Variablen belegt

      if ((hildesheimertitel ~ konstanzer_titel)) ||
          ((hildesheimerautor == konstanzerautor) &&
           (hildesheimerswort1 == konstanzerwort1) &&
           (hildesheimerswort2 == konstanzerwort2)) ||
          ((hildesheimerautor == konstanzerautor) &&
           (hildesheimerswort1 == konstanzerwort1) &&
           (hildesheimerswort3 == konstanzerwort3)) ||
          ((hildesheimerautor == konstanzerautor) &&
           (hildesheimerswort2 == konstanzerwort2) &&
           (hildesheimerswort3 == konstanzerwort3)) ||

          # wichtig, um Titel mit <= zwei Wörtern zu finden,
          # mindestens 3 von 4 Feldern und der Autor
          # müssen identisch sein:
          # so werden auch Titel erfasst, bei denen ein Wort
          # abweicht (z.B. eine Abkürzung oder irgendein
          # zusätzliches oder abweichendes Zeichen)
      {

```

```

((hidesheimerwort1 == konstanzerautor) &&
(hidesheimerwort2 == konstanzerwort1) &&
(hidesheimerwort3 == konstanzerwort2) &&
(hidesheimerwort4 == konstanzerwort3)))

# zum Auffinden der Titel ohne Autor
# im Konstanzer File

{
    print hidesheimertitel >> "end.txt"
}
}
close ("Ti_hi.txt")
}
}

```

```

BEGIN {print "Einzelne Notationen:"
      main()
}

function main()
{
    while (getline < "end_sort.txt" > 0 )    # alle Zeilen der sortierten Ergebnisdatei des Abgleichs auslesen
    {
        for (x=1 ; x<999 ; x++)              # von 1 bis 999 hochzählen
        {
            if (x<10) x="00"x                # für Notationen die kleiner 10 bzw. 100 sind die
            if (x<100 && x>=10) x="0"x        # entsprechenden Veränderungen vornehmen

            if ((index($0, "csc "x) > 0 ))    # CSC-Notationen heraussuchen
            {
                print $0 >> "csc_"x".txt"      # in Datei mit Namen der entsprechenden Notation schreiben
                close ("csc_"x".txt")
            }
        }
    }
}

```

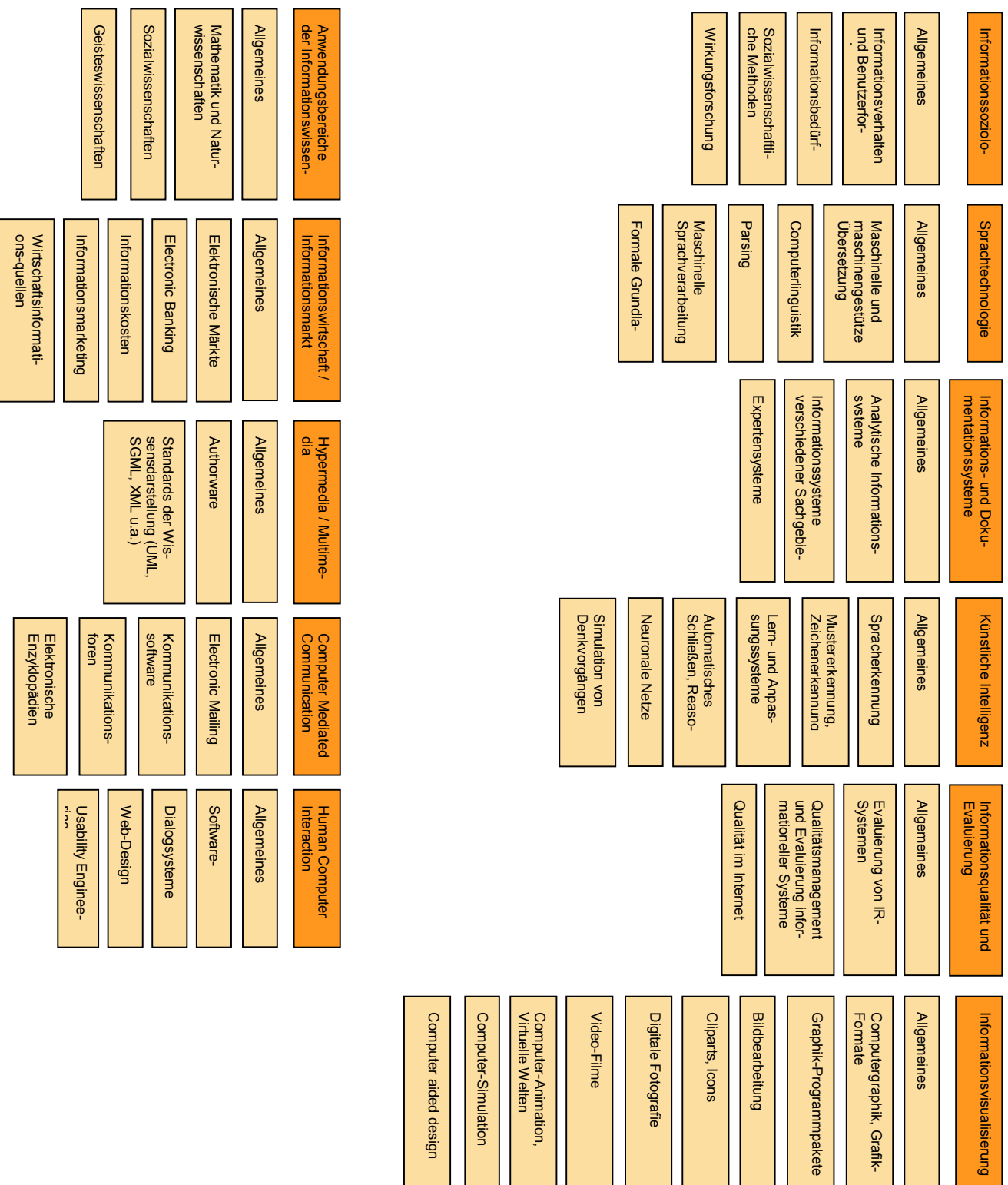
Allgemeines	Grundlagen / Theorie	Informationstechnik / Elektronische Datenverarbeitung	Informationsaufbereitung, -suche und -	Aspekte der Informationspraxis	Anwendungsbereiche	Sonstiges
Bibliografie	Thesauruskunde	Hardware	Information Retrieval	Informationsrecht	E-Commerce	
Biographien	Informationstheorie, Kybernetik	Rechenzentren	Informationsmanagement	Informationspolitik	Maschinelle Sprachverarbeitung	
Schriften und Berichte	Semantik	Betriebssysteme	Software-Ergonomie	Evaluierung und Qualitätsmanagement	Computeringuistik	
Lehrbücher	Methodologie	Dienstprogramme	Dialogsysteme	Informationskosten	Mathematik und Naturwissenschaften	
Sammelwerke	Terminologie	Programmierung	Elektronisches Publizieren	Informationsbedarf	Sozialwissenschaften	
Forschungseinrichtungen und -projekte	Klassifikation	Formale Sprachen	Hypertext, Hypermedia, Multimedia	Informationssoziologie	Geisteswissenschaften	
Institutionen, Organisationen und Verbände	Deskription	Datenerfassung	Informations- und Dokumentationssysteme	Informationspsychologie	Wirtschaft, Management, Organisation	
Beruf, Studium und Ausbildung	Kognitionstheorien	Datenorganisation	Bibliotheken		Künstliche Intelligenz	
Geschichte	Wissensrepräsentation	Datenbanken	Verlage		Informationsnetze	
Nachschlagewerke, Darstellung...	Indexieren	Datenschutz	Medien			
Zeitschriften	Katalogisierung	Fehlererkennung	Wissensvermittlung			
	Fuzzy-Logik	Datenfernverarbeitung	Informations- und Dokumentenstellen			
		Teilnehmersysteme				
		Datenkomprimierung				
		Internet				
		Firmen				
		Informations und Kommunikationstechnologie				

7 Anlage 15: Hildesheimer Klassifikation

## 6 Anlage 16: Hildesheimer Klassifikation (Teil

Allgemeines	Grundlagen / Theorie	Informationstechnik / Elektronische Datenverarbeitung	Information und Gesellschaft / Informationspolitik	Wissensvermittlung / Informations- und Dokumentationsstellen	Informationsmanagement / Wissensmanagement	Information Retrieval
Bibliographien	Allgemeines	Allgemeines	Allgemeines	Allgemeines	Allgemeines	Allgemeines
Biographien	Thesauruskunde	Hardware	Informationsrecht	Computerbasiertes Lernen und Unterrichten	Planung, Organisation	Retrievalsysteme und Dialogsprachen
Schriften und Berichte	Informationstheorie und -methodik Kybernetik	Betriebssysteme	Informationsethik	Digitale Bibliotheken	Dienstleistungen, Outsourcing	Modelle des Informations
Lehrbücher	Semantik, Semiotik	Dienstprogramme	Informationsfreiheit und - Förderungsprogramme	Elektronisches Publizieren	Innovationsforschung	Suchmaschinen, Suchmethoden im Internet
Sammelwerke	Dokumentationstypen	Programmierung		Verlage	Implementierungsforschung	Suchverfahren
Forschungseinrichtungen und -projekte	Terminologie	Datenerfassung- und -speicherung		Medien	Marketing, Informationsbroker	Software-Agenten
Institutionen, Organisationen und Verbände	Klassifikation	Datenverwaltung, EDV-Management		Bibliothekswesen/ Dokumentationswesen	Büroautomatisierung	Data Mining
	Deskription	Datenbanken		Dokumentenmanagementsysteme, Archivsysteme	Kosten, Nutzen, Leistung, Gebühren, Controlling	Knowledge Discovery
Beruf, Studium und Ausbildung	Informationspsychologie	Datensicherung			Public-Relations-Arbeit	
	Inhaltliche Erschließung	Fehlererkennung			Betriebliche Information & Kommunikation	
Geschichte	Fuzzy-Logik	Datenkomprimierung			Projektmanagement	
Nachschlagewerke, Darstellungen	Kommunikationstheorie und -modelle	Datennetze, Rechenzentren			Entscheidungssysteme	
Zeitschriften		Firmen				





#### 4 Anlage 17: Hildesheimer Klassifikation (Teil

Allgemeines	Grundlagen / Theorie	Informationstechnik / Elektronische Datenverarbeitung	Information Retrieval	Informationsqualität und Evaluierung	Informationsmanagement / Wissensmanagement	Informationswirtschaft / Informationsmarkt
Bibliographien	Allgemeines	Allgemeines	Allgemeines	Allgemeines	Allgemeines	Allgemeines
Schriften und Berichte	Informationstheorie	Hardware	Retrievalsysteme und Dialogsprachen	Evaluierung von IR-Systemen	Betriebliche Information & Kommunikation	Elektronische Märkte
Lehrbücher	Kommunikationstheorie und -modelle	Betriebssysteme	Modelle des Information Retrieval	Qualitätsmanagement und Evaluierung informationeller Systeme	Planung, Organisation	Electronic Banking
Sammelwerke	Fuzzy-Theorie	Dienstprogramme	Suchmaschinen, Suchmethoden im Internet	Qualität im Internet	Dienstleistungen, Outsourcing	Informationskosten
Nachschlagewerke, Darstellungen	Informationspsychologie	Formale Sprachen (ohne Programmiersprachen)	Suchverfahren		Innovationsforschung	Informationmarketing
Zeitschriften	Semantik, Semiotik	Datenerfassung und Speicherung	Software-Agenten		Implementierungsforschung	Information Brokering
Forschungseinrichtungen und -projekte	Terminologie	Datenverwaltung, EDV-Management	Data Mining, Knowledge Discovery		Marketing	Wirtschaftsinformationsquellen
Institutionen, Organisationen und Verbände	Thesauruskunde	Datenbanken			Büroautomatisierung	
Beruf, Studium und Ausbildung	Deskription	Datensicherung			Kosten, Nutzen, Leistung, Gebühren, Controlling	
	Inhaltliche Erschließung	Datenkomprimierung			Public-Relations-Arbeit	
	Klassifikation	Datenetze, Internet			Projektmanagement	
Geschichte	Dokumentationstypen	Rechenzentren			Entscheidungssysteme	
Biographien	Normen (UML, SGML, XML, u.a.)	Firmen			Internationales Informationsmanagement	

Human Computer Interaction	Sprachtechnologie	Computer Mediated Communication	Informationsvisualisierung	Wissensvermittlung / Informations- und Dokumentationsstellen	Informations- und Dokumentationsysteme
Allgemeines	Allgemeines	Allgemeines	Allgemeines	Allgemeines	Allgemeines
Software-Ergonomie	Maschinelle und maschinengestützte Übersetzung	Electronic Mailing	Computergraphik, Grafik-Formate	Computerbasiertes Lernen und Unterrichten	Analytische Informationssysteme
Dialogsysteme	Computeringuistik	Kommunikationssoftware	Graphik-Programmpakete	Digitale Bibliotheken	Informationssysteme verschiedener Sachgebiete
Web-Design	Parsing	Kommunikationsforen	Bildbearbeitung	Elektronisches Publizieren	
Usability			Digitale Fotografie, Digitaler Film	Elektronische Enzyklopädien	
Hypermedia/Multimedia	Maschinelle Sprachverarbeitung, Spracherkennung		Computer-Animation, Virtuelle Welten	Verlage	
	Formale Grundlagen		Computer-Simulation	Medien	
			Computer aided design	Bibliothekswesen, Dokumentationswesen	
Künstliche Intelligenz	Information und Gesellschaft / Informationspolitik	Informationswissenschaftliche Anwendungen in anderen Sachgebieten		Dokumentenmanagementsysteme, Archivalsysteme	
Allgemeines	Allgemeines	Allgemeines			
Mustererkennung, Zeichenerkennung	Informationsrecht	Mathematik und Naturwissenschaften			
	Informationsethik	Sozialwissenschaft			
Lernende Systeme und Anpassungssysteme (incl. Neuronale Netze)	Informationsgesellschaft	Geisteswissenschaft			
Automatisches Schließen, Reasoning	Informationsverhalten und Benutzerforschung				
Simulation von Denkvorgängen	Informationsbedürfnisse				
Diagnose	Wirkungsforschung				
Expertensysteme	Sozialwissenschaftliche Methoden				
	Förderungsprogramme				

### 3 Anlage 17: Hildesheimer Klassifikation (Teil

## 2 Anlage 18: Hildesheimer Klassifikation (Teil

Allgemeines	Grundlagen / Theorie	Informationstechnik / Elektronische Datenverarbeitung	Information Retrieval	Informationsqualität und Evaluierung	Informationsmanagement / Wissensmanagement	Informationswirtschaft / Informationsmarkt
Bibliographien	Informationstheorie	Hardware	Retrievalsysteme und Dialogsprachen	Evaluierung von IR-Systemen	Betriebliche Information & Kommunikation	Elektronische Märkte
Schriften und Berichte	Kommunikationstheorie und -modelle	Betriebssysteme	Modelle des Information Retrieval	Qualitätsmanagement und Evaluierung Informationeller Systeme	Planung, Organisation	Electronic Banking
Lehrbücher	Fuzzy-Theorie	Dienstprogramme	Suchmaschinen, Suchmethoden im Internet	Qualität im Internet	Dienstleistungen, Outsourcing	Informationskosten
Sammelwerke	Informationspsychologie	Programmierung	Suchverfahren	Sonstiges	Innovationsforschung	Informationsmarketing
Nachschlagewerke, Darstellungen	Semantik, Semiotik	Formale Sprachen (ohne Programmiersprachen)	Software-Agenten		Implementierungsforschung	Information Brokering
Zeitschriften	Terminologie	Datenerfassung und-speicherung	Data Mining, Knowledge Discovery		Marketing	Wirtschaftsinformationsquellen
Forschungseinrichtungen und -projekte	Thesauruskunde	Datenverwaltung, EDV-Management	Sonstiges		Büroautomatisierung	Sonstiges
Institutionen, Organisationen und Verbände	Description	Datenbanken			Kosten, Nutzen, Leistung, Gebühren, Controlling	
	Inhaltliche Erschließung	Datensicherung			Public-Relations-Arbeit	
	Klassifikation	Datenetze, Internet			Projektmanagement	
Beruf, Studium und Ausbildung	Dokumentationstypen	Rechenzentren			Entscheidungssysteme	
Geschichte	Normen (UML, SGML, XML u.a.)	Sonstiges			Internationales Informationsmanagement	
Biographien	Sonstiges				Sonstiges	
Firmen						

Human Computer Interaction	Sprachtechnologie	Computer Mediated Communication	Informationsvisualisierung	Wissensvermittlung / Informations- und Dokumentationsstellen	Informations- und Dokumentationsysteme
Software-Ergonomie	Maschinelle und maschinengestützte Übersetzung	Electronic Mailing	Computergraphik, Grafik-Formate	Computerbasiertes Lernen und Unterrichten	Analytische Informationssysteme
Dialogsysteme	Computerlinguistik	Kommunikationssoftware	Graphik-Programmpakete	Digitale Bibliotheken	Informationssysteme verschiedener Sachgebiete
Web-Design	Parsing	Kommunikationsforen	Bildbearbeitung	Elektronisches Publizieren	Sonstiges
Usability	Maschinelle Sprachverarbeitung / Spracherkennung	Sonstiges	Digitale Fotografie, Digitaler Film	Elektronische Enzyklopädien	
Hypermedia/Multimedia	Formale Grundlagen		Computer-Animation, Virtuelle Welten	Verlage	
Sonstiges	Sonstiges		Computer-Simulation	Medien	
			Computer aided design	Bibliothekswesen, Dokumentationswesen	
			Sonstiges	Dokumentenmanagementsysteme, Archivsysteme	
				Sonstiges	
Künstliche Intelligenz	Information und Gesellschaft / Informationspolitik	Informationswissenschaftliche Anwendungen in anderen Sachgebieten			
Mustererkennung, Zeichnerkennung	Informationsrecht	Mathematik und Naturwissenschaft-			
	Informationsethik	Sozialwissenschaft-			
Lernende Systeme und Anpassungssysteme (incl. Neuronale Netze)	Informationsgesellschaft	Geisteswissenschaft-			
Automatisches Schließen, Reasoning	Informationsverhalten und Benutzerforschung	Sonstiges			
Simulation von Denkvorgängen	Informationsbedürfnisse				
Diagnose	Wirkungsforschung				
Expertensysteme	Sozialwissenschaftliche Methoden				
Sonstiges	Förderungsprogramme				
	Sonstiges				

**Anlage 19:** Auszug aus der Hildesheimer informationswissenschaftlichen Klassifikation mit den zugeordneten Notationen

<b>7.2 Klassifikation</b>	<b>Anzahl der Titel</b>	<b>Zugehörige Notationen der bestehenden Klassifikation der Hildesheimer UB</b>
<b>7.3 Informationsvisualisierung</b>		
Computergraphik, Grafik-Formate	286	CSC 730, CSC 731, CSC 732, CSC 733, CSC 735, CSC 736, CSC 737, CSC 738, KUN 322,
Graphik-Programmpakete		
Bildbearbeitung	124	CSC 740, CSC 741, CSC 742, CSC 743, CSC 744, CSC 745, CSC 746, CSC 747, CSC 748, CSC 749
Digitale Fotografie, Digitaler Film	3	
Computer-Animation, virtuelle Welten	62	CSC 737,
Computer-Simulation	77	TFE 791, CSC 761, CSC 762, CSC 763, CSC 764, CSC 765, CSC 766, CSC 767, CSC 768, CSC 769,
Computer aided design	39	CSC 860, CSC 861, CSC 865
Sonstiges	6	

<b>7.4 Künstliche Intelligenz</b>		
Mustererkennung, Zeichenerkennung	140	CSC 750, CSC 751, CSC 752, CSC 753, CSC 754, CSC 755, CSC 758
Lernende Systeme und Anpassungssysteme (incl. Neuronale Netze)	135	CSC 726, MAT 726, CSC 726, CSC 751, LIN 480
Automatisches Schließen, Reasoning	124	CSC 723
Simulation von Denkvorgängen		
Diagnose		
Expertensysteme	134	BWL 687, REC 084, CSC 721
Sonstiges	361	CSC 720, CSC 721, CSC 722, CSC 724, CSC 725, CSC 728, CSC 729

# **Anlage 20 (Teil 1):**

## **Protokoll des Gesprches mit Edwin Pretz vom**

### **04.12.2001**

*Anwesende Personen:* Edwin Pretz (EDV-Beauftragter der UB Hildesheim), Peter Hanke

*Uhrzeit:* 11.00 Uhr - 12.00 Uhr

#### **Themen und Inhalte:**

- ⇒ Problematik: Wie kann man den Bestand der UB Hildesheim und anderer Standorte als Textdatei erhalten, um so ber einen Abgleich die informationswissenschaftlichen Bcher in Hildesheim zu finden?
- ⇒ Generelle Probleme dabei:
  - Bibliothekare scheuen sich, Gesamtbestnde oder auch Teilbestnde aus Themengebieten als Textdatei verfgbar zu machen (rechtliche Problematik). Ursache: Man mchte die eigene Arbeitsleistung vor fremdem Zugriff schtzen.
  - Hohe Belastung der Datenbank: im laufenden Betrieb nicht mglich
- ⇒ Mglichkeit zur Erfassung des Gesamtbestandes der UB Hildesheim (Pica-System): Datenbankabfrage ber SQL-Skript
  - wahrscheinlich sehr zeitaufwendig (Prfung durch Hr. Pretz soll erfolgen)
- ⇒ Mglichkeiten zur Erfassung des relevanten Bestandes der UB Hildesheim (ber Notationen, die Bezug zur Informationswissenschaft aufweisen):
  - ber kleineres Programm, welches eine Anfrage (z.B. nach Notationen) an die neue Such-Maske der UB schickt:
    - Problem der Ausgabe (immer nur begrenzte Anzahl von Titeln) und Abspeicherung (nur ber e-mail und auch hier nur Ausgabe einer begrenzten Anzahl von Titeln mglich)
  - Abfrage ber WinIWB-Oberflche (Clientsoftware, Arbeitsmaske der UB-Mitarbeiter)

→ Relevante Notationen werden ohne Trefferbeschränkung ausgegeben und können als Textdatei abgespeichert werden.



## **Anlage 20 (Teil 2):**

- Einarbeitung in WinIWB und die Regelung der Rechtevergabe könnte durch Hr. Pretz erfolgen.
- Problematik: Bücher, die ins Magazin kommen, verlieren ihre Notation, und bekommen eine neue Zahlenkombination ohne Buchstabenfolge.
- ⇒ Möglichkeiten zur Erfassung des relevanten Bestandes der anderen Standorte:
  - WinIWB-Oberfläche erlaubt durch Verbundsystem über Göttingen auch Zugriff auf RVK (SYBIS-Datenbank):
    - allerdings hier wahrscheinlich nur Abfrage über Titel, Autor, Signatur möglich und wohl nicht über Klassifikationsmerkmale
    - Problem einer hohen Belastung der Datenbank? (eine Frage des Umfanges der informationswissenschaftlichen Bestände in Regensburg und Konstanz)
  - Ob die WinIWB-Oberfläche auch einen Zugriff auf die Konstanzer UB zulässt, ist nicht klar, eventuell nur auf den SWB (Südwestdeutscher Bibliotheksverbund).
- ⇒ Möglichkeiten zum Abgleich der relevanten Bestände:
  - kleineres Programm, welches eine Anfrage (z.B. nach Autor, Titel, Signatur) an die neue Such-Maske der UB schickt:
    - Problem der Ausgabe (immer nur begrenzte Anzahl von Titeln) und Ab-speicherung (nur über e-mail und auch hier nur Ausgabe einer begrenzten Anzahl von Titeln)
  - Kombinierte Anfrage an Regensburg, Konstanz und Hildesheim über KVK (Karlsruher Virtueller Katalog) (eigenes Programm wäre zu schreiben)

## **Anlage 21 (Teil 1):**

### **Protokoll des Abschlußgespräches zur Magisterarbeit mit Ansprechpartnern der Bibliothek (vom 17.04.2002)**

*Anwesende Personen:* Maria Elisabeth Müller (Leiterin der UB Hildesheim), Edwin Pretz (EDV-Beauftragter der UB Hildesheim), Peter Hanke

*Zeit:* 10.30 Uhr – 12.00 Uhr

#### **Themen und Inhalte:**

- ⇒ Korrekte Bezeichnungen und Begriffe:
  - Der korrekte Begriff für eine Notation, die bei einem Titel neben der Signatur zusätzlich eingetragen wird, lautet Nebenstelle.
  - Konkordanz: völlige Deckungsgleichheit zwischen zwei Rubriken einer Klassifikation
  - Beschreibung der Rubriken: Klassifikationsbeschreibungen
  
- ⇒ Kritik von Hr. Pretz an der neuen informationswissenschaftlichen Klassifikation: Die Erstellung der Rubriken der Klassifikation erfolgte nicht nach dem Umfang der jeweils vorhandenen Titelmengen oder der Wichtigkeit für die Gesamtdisziplin, sondern pragmatisch nach anwendungsbezogenen Kriterien (Bezeichnungen der Lehrgebiete). Normalerweise wird beim Aufbau der Klassifikation darauf geachtet, dass Rubriken der gleichen Ebene die gleiche Gewichtung haben.
  
- ⇒ Evaluierung des Ergebnisses:
  - Wieviele der Bücher der neuen Klassifikation stehen falsch?
  - Eine umfangreichere Evaluierung der Arbeit müsste vor Übernahme in den OPAC erfolgen; der Wert für Studierende muss erwiesen werden, bevor umfangreichere Arbeiten zur Bereitstellung für Benutzer durchgeführt werden (Hr. Pretz).

- Vergleich der Treffermengen von gleichlautenden Rubriken der UB Hildesheim und der neuen Klassifikation: Welche Titel fehlen bei der Suchanfrage an UB und warum?

## Anlage 21 (Teil 2):

⇒ Änderungsvorschläge für die neue Klassifikation:

- „Allgemeines“ bezeichnet häufig überblicksartige Werke, die Wissen mehrerer Teilgebiete umfassen. Hier wäre „Sonstiges“ besser, da dort Werke erfasst werden sollen, die sonst nirgendwo untergebracht werden können (Fr. Müller).
- „Firmen“ passt besser unter „Allgemeines“, da es nicht nur den Bereich EDV betrifft (Begrifflichkeit hängt hier nicht am Fach, wie in den anderen Kategorien) (Hr. Pretz).
- Einzelne Kategorien dürfen mehr als zehn Teilgebiete enthalten (kein Problem für Einbau in bestehende Klassifikation der UB Hildesheim) (Hr. Pretz).

⇒ Erleichterungen für den Benutzer:

- Eine Klassifikationsbeschreibung mit Synonymen unter Erläuterungen der Themenbereiche muss erarbeitet werden (Fr. Müller).
- Eine Suchhilfe nach Themengebieten innerhalb der Systematik sollte dem Benutzer angeboten werden, damit man sich nicht erst mit der gesamten Klassifikation vertraut machen muss (Hr. Pretz).

⇒ Einbau der neuen Klassifikation in das Suchangebot:

- Problem für den Benutzer: Er müsste sich mit zwei Klassifikationen auseinandersetzen (Hr. Pretz).
- Problem der Einstellung neuer Bücher: Automatisierte Kopplung zwischen Notationen und Rubriken der Klassifikation ist nicht möglich (Es besteht keine Konkordanz.).
- Problem der Parallelität zu bestehender Klassifikation: Doppelstellen (Eindeutigkeit in Frage gestellt, verwirrend für Benutzer) (Hr. Pretz).
- Problem der weiteren Pflege (IKK-Klassifikation wurde nie weiter gepflegt)
- Vorschlag Hanke: Ermittelte Titel werden „statisch“ zur Verfügung gestellt, neue Titel erhalten als Nebenstelle eine Notation der Klassifikation

und können so per Datenbankabfrage dem Benutzer zugänglich gemacht werden.

### **Anlage 21 (Teil 3):**

- Vorschlag Pretz: Loslösung von bestehender Klassifikation, um keinen Eintrag in Datenbank machen zu müssen. Die beim Abgleich ermittelten Titel werden über ihre Signaturen den einzelnen Rubriken zugeordnet. Beim Anklicken einer Rubrik wird automatisch eine Suchanfrage nach den entsprechenden Signaturen gestartet. Neue Titel werden automatisiert eingestellt, indem bei bestimmten Suchanfragen die Signaturen der Treffer automatisch der Rubrik zugeordnet werden.
- Weiterer Vorschlag Pretz: Programm durchläuft Datenbank und ordnet Titel anhand der bestehenden Indexierung über Titel und Schlagwort auf neue Klassifikation zu. Dies wäre dann auch für andere Fachbereiche möglich.